PHYSIC

VOLUME 85

THE PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

Published by the American Institute of Physics for the

AMERICAN PHYSICAL SOCIETY

Officers

President

C. C. LAURITSEN
California Institute of Technology
Pasadena, California

Vice-President

J. H. VAN VLECK
Harvard University
Cambridge, Massachusetts

Secretary

K. K. DARROW
Bell Telephone Laboratories
Office:
Columbia University
New York 27, New York

Treasurer

G. B. PEGRAM
Columbia University
New York 27, New York

Local Secretary for the Pacific Coast

J. KAPLAN
University of California at Los Angeles
Los Angeles, California

Managing Editor

S. A. GOUDSMIT
Brookhaven National Laboratory
Upton, Long Island
New York

BOARD OF EDITORS

Editor

S. A. GOUDSMIT

Assistant Editor

S. PASTERNACK

Associate Editors

J. BARDEEN	HARVEY BROOKS	M. GELL-MANN
T. BONNER	R. F. CHRISTY	R. HOFSTADTER
E. G. HERB	C. L. CRITCHFIELD	A. W. LAWSON
W. E. LAMB, JR.	M. DEUTSCH	A. PAIS
A. J. F. SIEGERT	T. H. JOHNSON	R. SERBER
W. E. STEPHENS	BRUNO ROSSI	C. H. TOWNES

Manuscripts for publication should be submitted to S. A. Goudsmit, Brookhaven National Laboratory, Upton, Long Island, New York. The authors' institutions are requested to pay a publication charge of $3 per page which, if honored, entitles them to 100 free reprints. Instructions will be sent with galley proofs.

Proof and all correspondence concerning papers in the process of publication should be addressed to the Publication Manager, American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price

United States and Canada.....$15.00
Elsewhere.....$17.00

Back Numbers[1]

Yearly back number rate when complete year is available
Prior to 1949: $20.00
Thereafter: $27.00

Single copies: $2.00 each, prior to July, 1929;
$1.00 each, July, 1929–December, 1941;
$2.00 each, January, 1942–December, 1946;
$1.00 each, January, 1947–December, 1948;
$1.25 each, thereafter.

General index, 1893–1930: $4.00.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Changes of address in the case of members of the American Physical Society should be addressed to the Treasurer; in the case of other subscribers to the American Institute of Physics.

[1] For Series I inquire at Physics Department, Cornell University, Ithaca, New York.

The Physical Review is published semi-monthly at Prince and Lemon Streets, Lancaster, Pennsylvania.

Entered as second-class matter March 6, 1917, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879.

Acceptance for mailing at special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized January 27, 1925.

THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 85, NO. 2 JANUARY 15, 1952

Production Cross Sections for π^+- and π^--Mesons by 340-Mev Protons on Carbon and Lead at 90° to the Beam*

C. RICHMAN, M. WEISSBLUTH,† AND H. A. WILCOX‡
Radiation Laboratory, Department of Physics, University of California, Berkeley, California
(Received October 2, 1951)

Carbon and lead targets were bombarded by 340-Mev protons produced by the Berkeley 184-inch synchro-cyclotron. The spectrum of π^+- and π^+-mesons produced at 90°±12° to the beam was measured by the use of nuclear emulsions embedded in absorbers. The total cross section for the production of π^+-mesons from carbon at this angle is $(2.3\pm0.5)\cdot10^{-28}$ cm² ster^{-1}, and the π^+ to π^- ratio is 5.1±1.0. For lead the cross section for π^+-mesons at this angle is $(7\pm2.1)\cdot10^{-28}$ cm² ster^{-1} and the π^+ to π^- ratio is 1.5±0.4.

I. INTRODUCTION

THE artificial production of mesons was first achieved by E. Gardner and C. M. G. Lattes in 1948.[1] In their experiments they used the 380-Mev internal α-particle beam produced by the Berkeley 184-inch synchro-cyclotron. The experiment was carried out by bombarding a target on a probe inside the cyclotron. The charged mesons produced by the beam were bent away from the beam by the cyclotron magnet and were detected in suitably placed nuclear emulsions.

When the 184-inch cyclotron was converted so that protons could be accelerated to 340 Mev, and when subsequently these protons were deflected out of the cyclotron tank, the experimental possibilities for studying production of mesons were greatly improved. First, all the energy was concentrated in a single nucleon so that meson-production cross sections could be expected to be much greater. Second, it was of great help in these experiments to have the beam out of the vacuum tank and free of the cyclotron magnetic field even though one lost a factor of about a thousand in intensity in bringing the beam out.

The problem of the production of charged π-mesons from a light and a heavy nucleus by protons was undertaken. For the light nucleus carbon was chosen, and for the heavy nucleus lead was chosen. The method of detection was by means of nuclear emulsions which had been found to be very good for identifying mesons.

II. EXPERIMENTAL METHOD

The deflected proton beam from the cyclotron passes through a slit just outside of the cyclotron tank and then through a steering magnet. The beam then goes through a long evacuated pipe which passes through the main concrete shielding around the cyclotron. While passing through this shielding the beam is collimated once more. It then emerges into a concrete shielded pace known as the "cave" where the experiment is set up. The energy of the beam is 340±5 Mev. For this experiment the beam was collimated to a cross section at the target of about $1\frac{1}{4}\times1\frac{1}{4}$ square inches. The current was approximately 4×10^{-11} amp.

The arrangement of the apparatus in the cave is shown schematically in Fig. 1. The beam enters the cave through the exit window of the evacuated pipe. It then passes through the target under investigation. The targets used were graphite and lead rectangular foils 3 in.×5 in. and of various thicknesses. The beam continues in air and is finally caught in a Faraday cup of sufficient area and thickness to catch and stop all the protons.[2] The charge collected by the Faraday cup was accumulated on a condenser whose voltage was measured at convenient intervals. In this way the number of protons bombarding the target is known to ±5 percent and one can calculate the absolute cross

* Some of the results of this experiment were given at the Physical Society meeting at Stanford University in December, 1949, and Phys. Rev. **78**, 496 (1950).
† Now at Stanford University, Stanford, California.
‡ Now at the Naval Ordnance Test Station, China Lake, California.
[1] E. Gardner and C. M. G. Lattes, Science **107**, 270 (1948).
[2] The Faraday cup and associated circuits were kindly made available by Mr. V. Z. Peterson.

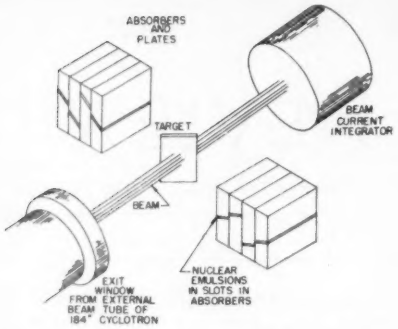


FIG. 1. Schematic drawing of the arrangement of the apparatus in the "cave."

sections. One of the advantages of having the beam out of the tank is the ease with which the beam current can be integrated.

When a nucleus like carbon or lead is bombarded by 340-Mev protons, one obtains at any angle to the beam a spectrum of π^+- and π^--mesons. To study such a spectrum the following scheme was used: The mesons emitted by the target are slowed down by absorbers which are placed at a given angle to the beam. Nuclear emulsions are embedded in the absorbers. The mesons are observed and identified in these emulsions at the ends of their tracks. Consequently the position of a meson in the emulsion reveals its energy, since it had to traverse a certain thickness of absorber material to reach the particular point in the emulsion. Furthermore a count of the number of mesons stopped in a certain volume of emulsion can be directly correlated to the number of mesons of a certain energy interval emitted into a certain solid angle by the target.

Figure 1 shows the absorber blocks with the embedded nuclear plates that make up the detectors. In this experiment the absorbers were located at 90° to the beam. The absorber blocks were 3 in.×3.5 in. ×0.982 in. with a 15° cut to hold the plates. A piece of black masking tape around the cut kept the plates from being light struck. The absorber block nearest the target was at a distance of 7.15 cm from the center of the target. It was made of aluminum and stopped mesons of energies from 0 to 37 Mev. The next absorber block which was placed immediately behind the first was made out of copper and was capable of stopping mesons from 37 to 85 Mev. The third absorber was also made of copper and covered the range between 85 and 125 Mev. The absorbers were large enough so that the scanned portions of the photographic plates were effectively surrounded by an infinite sea of absorber medium. Thus except near the extreme edges, there was no significant loss of particles from the absorber due to multiple scattering.

The photographic plates used were Ilford Nuclear Research plates, type C-3, with an emulsion thickness of 100 microns. These plates are sufficiently sensitive to render visible the entire track of a μ^+-meson originating from the decay of a π^+-meson in the emulsion; the plates are not so sensitive, however, as to obliterate the characteristic increase in ionization when a meson approaches the end of its range. Fast electrons are not observed in this type of emulsion.

In principle, it would be desirable to expose the photographic plates for a sufficiently long time to give a high density of mesons and make the counting very rapid. This, however, cannot be done in practice. The cross section for the scattering of protons by the target nuclei is of the order of a thousand times larger than the meson-production cross section. Consequently the duration of the exposure necessary to produce a plate such that a meson can be reliably recognized amidst other tracks is entirely controlled by the background radiation.

The background consists, first, of all charged particles which come directly from the target, and second, from the ambient flux of radiation in the cave which arises from collimating and stopping the proton beam. The background from the target is unavoidable in the present arrangement. The background due to the ambient radiation is reduced by lead shielding (not shown in Fig. 1) placed around the absorbers which hold the photographic plates. These two sources of background lead to a qualitative difference in the appearance of a plate which records low energy mesons from one which records high energy mesons. At 90° to the beam, the low energy plates show background tracks with a preferred orientation, namely pointing to the target. On the other hand, the plates exposed in a position to record the high energy mesons show background tracks oriented in all directions; which indicates that here the charged particles do not come directly from the target. These charged particles are produced by the ambient neutron flux.

The experimental arrangement that we have described works well for studies of meson production at angles of 90° and larger with respect to the direction of the beam. As one goes from 90° forward, that is, toward the direction of the beam, however, one finds that the background of scattered particles increases very rapidly; this makes the scanning of the plates difficult except for very light exposures. For studying the production of mesons at the smaller angles it is best to use some method of reducing the direct background from the target, such as bending the mesons away from the beam in a channel which is placed inside of a magnetic field.[3]

III. CLASSIFICATION OF MESONS

The well-known characteristics of mesons in emulsions enable one to find the mesons and very often to

[3] Cartwright, Richman, Whitehead, and Wilcox, Phys. Rev. **78**, 823 (1950).

say what kind of mesons they are. The facts are as follows: A positive π-meson which comes to the end of its range decays into a positive μ-meson and another light neutral particle which is presumably a neutrino.[4] On the other hand, negative π-mesons which come to rest in the emulsion generally produce nuclear stars.[5] It has been found from studies[6] on the prong distribution of stars from magnetically sorted mesons that 73 ± 2 percent of the negative π's produce a nuclear star of one or more observable prongs in emulsions of the type used here. The remaining 27 percent of the π^- mesons end in the emulsion without initiating an observable nuclear star; in these cases, presumably, neutrons or photons are emitted. Thus it is seen that positive and negative π-mesons have characteristics which render them distinguishable when appearing together in an emulsion.

We now consider a more detailed classification based on the actual appearance of mesons in the emulsion:

1. A meson ends in a star of two or more prongs or of a single heavy prong. This is identified as a π^--meson. No stars have been observed from π^+- or μ^+-mesons. It has been shown[7] that 8.7 ± 1.7 of the μ^--mesons stopped in emulsions form stars. This gives a negligible correction in our experiment.

2. A meson ends with the emission of another meson. The secondary meson is identified as such by virtue of the characteristic small angle scattering and the rapid increase in grain density near the end of its track. Furthermore, the track is about 600 microns long which is the range in emulsion of a μ-meson originating from a stationary π-μ decay. This is then a π-μ decay. A π-μ decay is counted as a π^+-meson.

3. Mesons appear in the emulsion and stop with no observable subsequent events. Such a meson, designated as a ρ-meson may be one of the following types: (a) It may be a μ^+-meson originating from a π^+-meson which has come to rest in the glass or absorber surrounding the emulsion. Such mesons need not be counted. (b) It may be a negative π-meson which does not initiate a star with observable prongs. This occurs, as mentioned above, for 27 percent of the stopped π^--mesons. Such mesons are taken into account by the correction applied to the number of definitely identified negative π-mesons. (c) The ρ-meson may be a positive or negative μ from a positive or negative π which decayed in flight. An over-all correction for decay in flight is applied and will be discussed in Sec. IV.

4. A meson stops in the emulsion with the emission of a particle which makes a thin, lightly ionizing track. The secondary track leaves the emulsion in a distance too short to make it possible to identify the particle which produced it. It is clear that this is either a π-μ decay or it is a π^--meson which produced a one-prong lightly ionizing star. To decide on the classification of these mesons, we go back to some data which has been obtained by F. L. Adelman and S. B. Jones with magnetically sorted π^--mesons. They have found that out of 65 one-prong stars produced by π^-'s, 14 had fast prongs and were confusable with π-μ endings. However, out of these 14 cases, 12 had short heavy recoils which we call "clubs." It is known that π-μ's do not exhibit such clubs. Only 2 out of the 65 π^- mesons did not exhibit such clubs and would then be confusable with π-μ events. So by examining these cases for "clubs" we will make less than a 1 percent error in the total number of π^--mesons, and it turns out even less of an error in the total number of π^+-mesons.

IV. CALCULATION OF CROSS SECTION

Let a total of n protons be incident on a target having m nuclei per cm^2. Let the cross section for the production of mesons per unit solid angle at the angle θ to the beam and per unit energy interval at the energy E be denoted by $d\sigma/d\Omega dE$. The mesons are observed by scanning a certain volume of emulsion embedded in the absorber. Let a given volume element be at a distance R from the target and be at such a depth in the absorber to stop mesons around the energy E. It is easy to show that if one observes N mesons per unit volume of emulsion then

$$\frac{d\sigma}{dEd\Omega}=\frac{NR^2}{nm(dE/dx)_{em}}, \tag{1}$$

where $(dE/dx)_{em}$ is the energy loss per unit distance of mesons of energy E in emulsion.

The emulsions were scanned with a Spencer binocular microscope using a magnification of about 600×. The microscopes were equipped with a special stage designed and built by Mr. W. M. Brower of the Physics Department machine shop which enable an observer to reset the scope within about two microns.

Since photographic emulsions do not preserve their thicknesses upon development it was necessary to devise a technique to determine the thicknesses prior to development. This was done in the following way: The nuclear plates to be used in the meson experiment were exposed to a very weak beam of 380-Mev α-particles which entered the plates at an accurately known angle of about 20° to the emulsion surface. The α-particle tracks are then found after the plates have been used in the experiment and developed. By measuring the projected length, L, of such a track in the plane of the emulsion the original thickness was easily calculated. In this way emulsion thicknesses can be measured to an accuracy of about 3 percent.

[4] Lattes, Muirhead, Occhialini, and Powell, Nature **159**, 694 (1947); Lattes, Occhialini, and Powell, Nature **160**, 453, 486 (1947); Burfening, Gardner, and Lattes, Phys. Rev. **75**, 382 (1949).
[5] D. H. Perkins, Nature **159**, 126 (1947); Lattes, Occhialini, and Powell, Nature **160**, 453, 486 (1947).
[6] F. L. Adelman and S. B. Jones, Phys. Rev. **75**, 1468 (A) (1949).
[7] E. P. George and J. Evans, Proc. Phys. Soc. (London) **64**, 193 (1951).

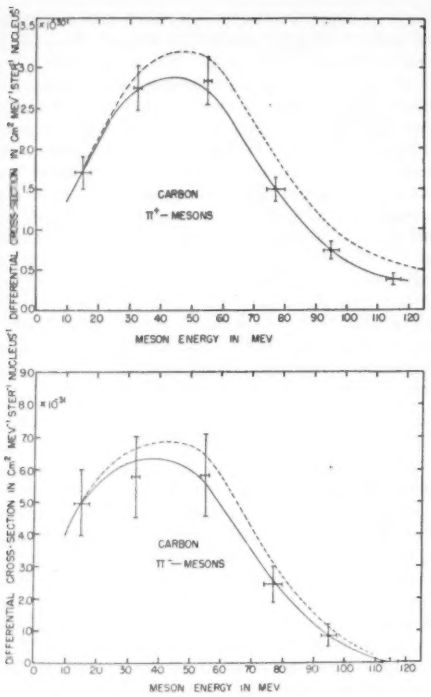


FIG. 2. Differential cross sections for the production of π^+- and π^--mesons by 340-Mev protons on carbon. Angle of observation $=90°\pm12°$ to the beam direction. The dashed curves are the experimental curves corrected for nuclear absorption of the mesons.

There are two corrections which need to be made to the cross section as calculated by Eq. (1): First, mesons produced in the target will decay during the time it takes them to go from the target to the end of their ranges in the emulsion. The mean life of the π^+-meson is $2.54\pm0.12\times10^{-8}$ sec.[8] The mean life of the π^--meson is $2.92\pm0.32\times10^{-8}$ sec.[9] The proper time which has elapsed from the time the meson is created until it comes to rest is given by $\int[(1-\beta^2)^{\frac{1}{2}}dx/\beta c]$ where β is the velocity of the meson divided by c, the velocity of light, dx is the element of path length and the integration is carried out over the entire path of the meson. This correction is never more than 4 percent of the number of observed mesons. Second, the method of detection used in this experiment, namely, slowing the mesons down until they stop in the emulsion introduces an appreciable error in the cross section due to nuclear absorption in the stopping material. To show the effect of this correction the absorption cross section has been taken to be equal to the geometrical cross section of the nucleus independent of the meson energy. The corrected spectra are given by the dashed curves. The effect of the nuclear scattering of the mesons would be to shift the whole spectrum to slightly lower energies. No corrections have been made for this effect.

RESULTS AND CONCLUSIONS

The differential cross section in cm² per Mev per steradian per nucleus for the production of π^+- and π^--mesons from carbon at $90°\pm12°$ is shown in Fig. 2. The probable errors shown on the graph are the purely statistical errors. These are valid for the relative values of the cross sections. The whole curve is subject to an added uncertainly of about ±15 percent in the absolute values of the cross sections. The dashed curves are the experimental curves corrected for the absorption of mesons assuming that the cross section for absorption is equal to the nuclear area independent of meson energy.

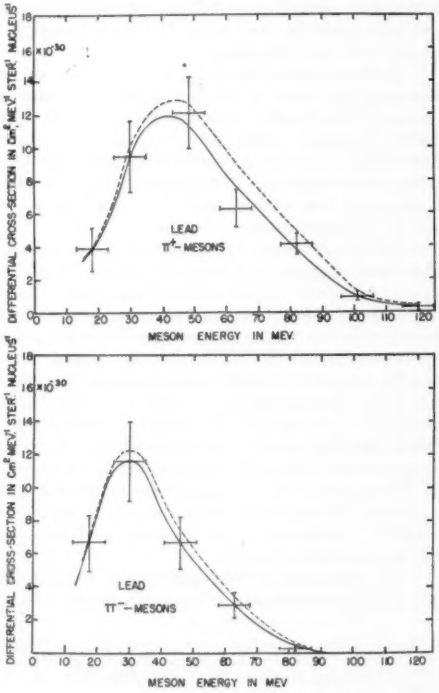


FIG. 3. Differential cross sections for the production of π^+- and π^--mesons by 340-Mev protons on lead. Angle of observation $=90°\pm12°$ to the beam direction. The dashed curves are the experimental curves corrected for nuclear absorption of the mesons.

[8] Chamberlain, Mozeley, Steinberger, and Wiegand, Phys. Rev. **79**, 394 (1950).
[9] Lederman, Booth, Byfield, and Kessler, Phys. Rev. **83**, 685 (1951).

The integral over energy of the π^+ spectrum (including the absorption correction) is $(2.3\pm0.5)\times10^{-28}$ cm^2 ster^{-1}. The integral over energy of the π^- spectrum is $(4.5\pm1.3)\times10^{-29}$ cm^2 ster^{-1}. This gives an average ratio of π^+- to π^--mesons at this angle of 5.1 ± 1.0. The uncertainty in this ratio is the statistical error.

Figure 3 shows the equivalent curves for lead. The integral over energy of the π^+ spectrum for this nucleus is $(7\pm2.1)\times10^{-28}$ cm^2 ster^{-1}, and the integral over energy of the π^- spectrum is $(4.7\pm1.9)\times10^{-28}$ cm^2 ster^{-1}. An average positive to negative ratio of 1.5 ± 0.4 is then obtained.

The decrease in the average π^+ to π^- ratio as one goes from carbon to lead is to be expected since the ratio of neutrons to protons is much greater for lead than for carbon. It is interesting to look at the π^+ to π^- ratio as a function of the energy of the mesons. For carbon the total number of π^--mesons found is low and so it is not clear what is the behavior of this ratio except that there is some indication that it increases as a function of the energy. In the case of lead this behavior becomes clearer. Figure 4 shows a plot of the ratio of π^-- to π^+-mesons from lead as a function of meson energy, and here one sees a definite decrease in this ratio at the higher energies. It has been pointed out by Chew and Steinberger[10] that when π-mesons are produced in the collision of protons with complex nuclei, π^+-mesons would be favored over π^--mesons because of the Pauli exclusion principle. The opposite would be true if the bombarding particle were a neutron. Furthermore, for a given proton energy the π^+ to π^- ratio should increase with meson energy. The data from carbon and lead supports this conclusion. They also conclude that for a given meson energy, the π^+ to π^- ratio should decrease as the proton energy is increased. Recently Block, Passman, and Havens[11] have studied the production of π^+- and π^--mesons from carbon by 381-Mev protons at 90° to the beam. They find, contrary to the above ideas, that the average π^+ to π^- ratio has increased to 11 ± 3. This increase has come about in the following way: the π^+ production cross section has increased by about a factor of 2, but the π^- production cross section has not changed very much. This seems to indicate that the elementary processes

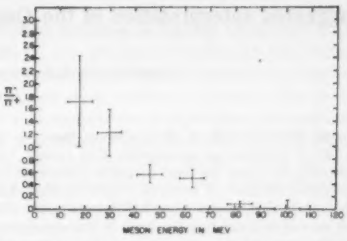


FIG. 4. The ratio of π^-- to π^+-mesons from lead at $90°\pm12°$ to the beam as a function of meson energy.

of the production of a π^+- and a π^--meson are perhaps of different character.

Henley and Huddlestone[12] have treated the production of π^+-mesons by protons on carbon from a phenomenological point of view. They have assumed that the mesons are produced in nucleon-nucleon collisions, and have used the measured cross sections of the production of π^+-mesons in proton-proton collisions without the contribution from deuteron formation.[2,13] The resulting spectrum of mesons at 90° to the beam depends rather critically on the momentum distribution that one takes for the nucleons. The theoretical problem has also been considered by Passman, Block, and Havens[14] from a similar point of view. They have used a Fermi degenerate gas model for the target nucleons, and have assumed that the final state nucleon-nucleon interaction is such that the π^+-meson always comes off with the maximum available kinetic energy. The agreement with the π^+ experimental data at 343 Mev and at 381 Mev turns out quite satisfactorily.

ACKNOWLEDGMENT

We wish to thank Professors E. O. Lawrence and R. L. Thornton for their interest and encouragement. We would like to thank Mr. John Foster for his generous help in the earlier stages of the work, and Mr. James Vale and the cyclotron crew for making the bombardments. This work was sponsored by the AEC.

[10] G. F. Chew and J. L. Steinberger, Phys. Rev. **78**, 497 (1950).
[11] Block, Passman, and Havens, Phys. Rev. **83**, 167 (1951).
[12] E. M. Henley and R. H. Huddlestone, Phys. Rev. **82**, 754 (1951).
[13] M. N. Whitehead and C. Richman, Phys. Rev. **83**, 855 (1951).
[14] Passman, Block, and Havens, Phys. Rev. **83**, 167 (1951).

A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I

DAVID BOHM*
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received July 5, 1951)

The usual interpretation of the quantum theory is self-consistent, but it involves an assumption that cannot be tested experimentally, *viz.*, that the most complete possible specification of an individual system is in terms of a wave function that determines only probable results of actual measurement processes. The only way of investigating the truth of this assumption is by trying to find some other interpretation of the quantum theory in terms of at present "hidden" variables, which in principle determine the precise behavior of an individual system, but which are in practice averaged over in measurements of the types that can now be carried out. In this paper and in a subsequent paper, an interpretation of the quantum theory in terms of just such "hidden" variables is suggested. It is shown that as long as the mathematical theory retains its present general form, this suggested interpretation leads to precisely the same results for all physical processes as does the usual interpretation. Nevertheless, the suggested interpretation provides a broader conceptual framework than the usual interpretation, because it makes possible a precise and continuous description of all processes, even at the quantum level. This broader conceptual framework allows more general mathematical formulations of the theory than those allowed by the usual interpretation. Now, the usual mathematical formulation seems to lead to insoluble difficulties when it is extrapolated into the domain of distances of the order of 10^{-13} cm or less. It is therefore entirely possible that the interpretation suggested here may be needed for the resolution of these difficulties. In any case, the mere possibility of such an interpretation proves that it is not necessary for us to give up a precise, rational, and objective description of individual systems at a quantum level of accuracy.

1. INTRODUCTION

THE usual interpretation of the quantum theory is based on an assumption having very far-reaching implications, *viz.*, that the physical state of an individual system is completely specified by a wave function that determines only the probabilities of actual results that can be obtained in a statistical ensemble of similar experiments. This assumption has been the object of severe criticisms, notably on the part of Einstein, who has always believed that, even at the quantum level, there must exist precisely definable elements or dynamical variables determining (as in classical physics) the actual behavior of each individual system, and not merely its probable behavior. Since these elements or variables are not now included in the quantum theory and have not yet been detected experimentally, Einstein has always regarded the present form of the quantum theory as incomplete, although he admits its internal consistency.[1–5]

Most physicists have felt that objections such as those raised by Einstein are not relevant, first, because the present form of the quantum theory with its usual probability interpretation is in excellent agreement with an extremely wide range of experiments, at least in the domain of distances[6] larger than 10^{-13} cm, and, secondly, because no consistent alternative interpretations have as yet been suggested. The purpose of this paper (and of a subsequent paper hereafter denoted by II) is, however, to suggest just such an alternative interpretation. In contrast to the usual interpretation, this alternative interpretation permits us to conceive of each individual system as being in a precisely definable state, whose changes with time are determined by definite laws, analogous to (but not identical with) the classical equations of motion. Quantum-mechanical probabilities are regarded (like their counterparts in classical statistical mechanics) as only a practical necessity and not as a manifestation of an inherent lack of complete determination in the properties of matter at the quantum level. As long as the present general form of Schroedinger's equation is retained, the physical results obtained with our suggested alternative interpretation are precisely the same as those obtained with the usual interpretation. We shall see, however, that our alternative interpretation permits modifications of the mathematical formulation which could not even be described in terms of the usual interpretation. Moreover, the modifications can quite easily be formulated in such a way that their effects are insignificant in the atomic domain, where the present quantum theory is in such good agreement with experiment, but of crucial importance in the domain of dimensions of the order of 10^{-13} cm, where, as we have seen, the present theory is totally inadequate. It is thus entirely possible that some of the modifications describable in terms of our suggested alternative interpretation, but

* Now at Universidade de São Paulo, Faculdade de Filosofia, Ciencias, e Letras, São Paulo, Brasil.

[1] Einstein, Podolsky, and Rosen, Phys. Rev. **47**, 777 (1933).

[2] D. Bohm, *Quantum Theory* (Prentice-Hall, Inc., New York, 1951), see p. 611.

[3] N. Bohr, Phys. Rev. **48**, 696 (1935).

[4] W. Furry, Phys. Rev. **49**, 393, 476 (1936).

[5] Paul Arthur Schilp, editor, *Albert Einstein, Philosopher-Scientist* (Library of Living Philosophers, Evanston, Illinois, 1949). This book contains a thorough summary of the entire controversy.

[6] At distances of the order of 10^{-13} cm or smaller and for times of the order of this distance divided by the velocity of light or smaller, present theories become so inadequate that it is generally believed that they are probably not applicable, except perhaps in a very crude sense. Thus, it is generally expected that in connection with phenomena associated with this so-called "fundamental length," a totally new theory will probably be needed. It is hoped that this theory could not only deal precisely with such processes as meson production and scattering of elementary particles, but that it would also systematically predict the masses, charges, spins, etc., of the large number of so-called "elementary" particles that have already been found, as well as those of new particles which might be found in the future.

not in terms of the usual interpretation, may be needed for a more thorough understanding of phenomena associated with very small distances. We shall not, however, actually develop such modifications in any detail in these papers.

After this article was completed, the author's attention was called to similar proposals for an alternative interpretation of the quantum theory made by de Broglie[7] in 1926, but later given up by him partly as a result of certain criticisms made by Pauli[8] and partly because of additional objections raised by de Broglie[7] himself.† As we shall show in Appendix B of Paper II, however, all of the objections of de Broglie and Pauli could have been met if only de Broglie had carried his ideas to their logical conclusion. The essential new step in doing this is to apply our interpretation in the theory of the measurement process itself as well as in the description of the observed system. Such a development of the theory of measurements is given in Paper II,[9] where it will be shown in detail that our interpretation leads to precisely the same results for all experiments as are obtained with the usual interpretation. The foundation for doing this is laid in Paper I, where we develop the basis of our interpretation, contrast it with the usual interpretation, and apply it to a few simple examples, in order to illustrate the principles involved.

2. THE USUAL PHYSICAL INTERPRETATION OF THE QUANTUM THEORY

The usual physical interpretation of the quantum theory centers around the uncertainty principle. Now, the uncertainty principle can be derived in two different ways. First, we may start with the assumption already criticized by Einstein,[1] namely, that a wave function that determines only probabilities of actual experimental results nevertheless provides the most complete possible specification of the so-called "quantum state" of an individual system. With the aid of this assumption and with the aid of the de Broglie relation, $\mathbf{p}=\hbar\mathbf{k}$, where $\mathbf{k}$ is the wave number associated with a particular fourier component of the wave function, the uncertainty principle is readily deduced.[10] From this derivation, we are led to interpret the uncertainty principle as an inherent and irreducible limitation on the precision with which it is correct for us even to conceive of momentum and position as simultaneously defined quantities. For if, as is done in the usual interpretation of the quantum theory, the wave intensity is assumed to determine only the probability of a given position, and if the $\mathbf{k}$th Fourier component of the wave function is assumed to determine only the probability of a corresponding momentum, $\mathbf{p}=\hbar\mathbf{k}$, then it becomes a contradiction in terms to ask for a state in which momentum and position are simultaneously and precisely defined.

A second possible derivation of the uncertainty principle is based on a theoretical analysis of the processes with the aid of which physically significant quantities such as momentum and position can be measured. In such an analysis, one finds that because the measuring apparatus interacts with the observed system by means of indivisible quanta, there will always be an irreducible disturbance of some observed property of the system. If the precise effects of this disturbance could be predicted or controlled, then one could correct for these effects, and thus one could still in principle obtain simultaneous measurements of momentum and position, having unlimited precision. But if one could do this, then the uncertainty principle would be violated. The uncertainty principle is, as we have seen, however, a necessary consequence of the assumption that the wave function and its probability interpretation provide the most complete possible specification of the state of an individual system. In order to avoid the possibility of a contradiction with this assumption, Bohr[3,5,10,11] and others have suggested an additional assumption, namely, that the process of transfer of a single quantum from observed system to measuring apparatus is inherently unpredictable, uncontrollable, and not subject to a detailed rational analysis or description. With the aid of this assumption, one can show[10] that the same uncertainty principle that is deduced from the wave function and its probability interpretation is also obtained as an inherent and unavoidable limitation on the precision of all possible measurements. Thus, one is able to obtain a set of assumptions, which permit a self-consistent formulation of the usual interpretation of the quantum theory.

The above point of view has been given its most consistent and systematic expression by Bohr,[3,5,10] in terms of the "principle of complementarity." In formulating this principle, Bohr suggests that at the atomic level we must renounce our hitherto successful practice of conceiving of an individual system as a unified and precisely definable whole, all of whose aspects are, in a manner of speaking, simultaneously and

[7] L. de Broglie, *An Introduction to the Study of Wave Mechanics* (E. P. Dutton and Company, Inc., New York, 1930), see Chapters 6, 9, and 10. See also Compt. rend. **183**, 447 (1926); **184**, 273 (1927); **185**, 380 (1927).

[8] *Reports on the Solvay Congress* (Gauthiers-Villars et Cie., Paris, 1928), see p. 280.

† *Note added in proof.*—Madelung has also proposed a similar interpretation of the quantum theory, but like de Broglie he did not carry this interpretation to a logical conclusion. See E. Madelung, Z. f. Physik **40**, 332 (1926), also G. Temple, *Introduction to Quantum Theory* (London, 1931).

[9] In Paper II, Sec. 9, we also discuss von Neumann's proof [see J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Verlag, Julius Springer, Berlin, 1932)] that quantum theory cannot be understood in terms of a statistical distribution of "hidden" causal parameters. We shall show that his conclusions do not apply to our interpretation, because he implicitly assumes that the hidden parameters must be associated only with the observed system, whereas, as will become evident in these papers, our interpretation requires that the hidden parameters shall also be associated with the measuring apparatus.

[10] See reference 2, Chapter 5.

[11] N. Bohr, *Atomic Theory and the Description of Nature* (Cambridge University Press, London, 1934).

unambiguously accessible to our conceptual gaze. Such a system of concepts, which is sometimes called a "model," need not be restricted to pictures, but may also include, for example, mathematical concepts, as long as these are supposed to be in a precise (i.e., one-to-one) correspondence with the objects that are being described. The principle of complementarity requires us, however, to renounce even mathematical models. Thus, in Bohr's point of view, the wave function is in no sense a conceptual model of an individual system, since it is not in a precise (one-to-one) correspondence with the behavior of this system, but only in a statistical correspondence.

In place of a precisely defined conceptual model, the principle of complementarity states that we are restricted to complementarity pairs of inherently imprecisely defined concepts, such as position and momentum, particle and wave, etc. The maximum degree of precision of definition of either member of such a pair is reciprocally related to that of the opposite member. This need for an inherent lack of complete precision can be understood in two ways. First, it can be regarded as a consequence of the fact that the experimental apparatus needed for a precise measurement of one member of a complementary pair of variables must always be such as to preclude the possibility of a simultaneous and precise measurement of the other member. Secondly, the assumption that an individual system is completely specified by the wave function and its probability interpretation implies a corresponding unavoidable lack of precision in the very conceptual structure, with the aid of which we can think about and describe the behavior of the system.

It is only at the classical level that we can correctly neglect the inherent lack of precision in all of our conceptual models; for here, the incomplete determination of physical properties implied by the uncertainty principle produces effects that are too small to be of practical significance. Our ability to describe classical systems in terms of precisely definable models is, however, an integral part of the usual interpretation of the theory. For without such models, we would have no way to describe, or even to think of, the result of an observation, which is of course always finally carried out at a classical level of accuracy. If the relationships of a given set of classically describable phenomena depend significantly on the essentially quantum-mechanical properties of matter, however, then the principle of complementarity states that no single model is possible which could provide a precise and rational analysis of the connections between these phenomena. In such a case, we are not supposed, for example, to attempt to describe in detail how future phenomena arise out of past phenomena. Instead, we should simply accept without further analysis the fact that future phenomena do in fact somehow manage to be produced, in a way that is, however, necessarily beyond the possibility of a detailed description. The only aim of a mathematical theory is then to predict the statistical relations, if any, connecting these phenomena.

3. CRITICISM OF THE USUAL INTERPRETATION OF THE QUANTUM THEORY

The usual interpretation of the quantum theory can be criticized on many grounds.[8] In this paper, however, we shall stress only the fact that it requires us to give up the possibility of even conceiving precisely what might determine the behavior of an individual system at the quantum level, without providing adequate proof that such a renunciation is necessary.[9] The usual interpretation is admittedly consistent; but the mere demonstration of such consistency does not exclude the possibility of other equally consistent interpretations, which would involve additional elements or parameters permitting a detailed causal and continuous description of all processes, and not requiring us to forego the possibility of conceiving the quantum level in precise terms. From the point of view of the usual interpretation, these additional elements or parameters could be called "hidden" variables. As a matter of fact, whenever we have previously had recourse to statistical theories, we have always ultimately found that the laws governing the individual members of a statistical ensemble could be expressed in terms of just such hidden variables. For example, from the point of view of macroscopic physics, the coordinates and momenta of individual atoms are hidden variables, which in a large scale system manifest themselves only as statistical averages. Perhaps then, our present quantum-mechanical averages are similarly a manifestation of hidden variables, which have not, however, yet been detected directly.

Now it may be asked why these hidden variables should have so long remained undetected. To answer this question, it is helpful to consider as an analogy the early forms of the atomic theory, in which the existence of atoms was postulated in order to explain certain large-scale effects, such as the laws of chemical combination, the gas laws, etc. On the other hand, these same effects could also be described directly in terms of existing macrophysical concepts (such as pressure, volume, temperature, mass, etc.); and a correct description in these terms did not require any reference to atoms. Ultimately, however, effects were found which contradicted the predictions obtained by extrapolating certain purely macrophysical theories to the domain of the very small, and which could be understood correctly in terms of the assumption that matter is composed of atoms. Similarly, we suggest that if there are hidden variables underlying the present quantum theory, it is quite likely that in the atomic domain, they will lead to effects that can also be described adequately in the terms of the usual quantum-mechanical concepts; while in a domain associated with much smaller dimensions, such as the level associated with the "fundamental length" of the order of 10^{-13} cm, the hidden variables

may lead to completely new effects not consistent with the extrapolation of the present quantum theory down to this level.

If, as is certainly entirely possible, these hidden variables are actually needed for a correct description at small distances, we could easily be kept on the wrong track for a long time by restricting ourselves to the usual interpretation of the quantum theory, which excludes such hidden variables as a matter of principle. It is therefore very important for us to investigate our reasons for supposing that the usual physical interpretation is likely to be the correct one. To this end, we shall begin by repeating the two mutually consistent assumptions on which the usual interpretation is based (see Sec. 2):

(1) The wave function with its probability interpretation determines the most complete possible specification of the state of an individual system.

(2) The process of transfer of a single quantum from observed system to measuring apparatus is inherently unpredictable, uncontrollable, and unanalyzable.

Let us now inquire into the question of whether there are any experiments that could conceivably provide a test for these assumptions. It is often stated in connection with this problem that the mathematical apparatus of the quantum theory and its physical interpretation form a consistent whole and that this combined system of mathematical apparatus and physical interpretation is tested adequately by the extremely wide range of experiments that are in agreement with predictions obtained by using this system. If assumptions (1) and (2) implied a unique mathematical formulation, then such a conclusion would be valid, because experimental predictions could then be found which, if contradicted, would clearly indicate that these assumptions were wrong. Although assumptions (1) and (2) do limit the possible forms of the mathematical theory, they do not limit these forms sufficiently to make possible a unique set of predictions that could in principle permit such an experimental test. Thus, one can contemplate practically arbitrary changes in the Hamiltonian operator, including, for example, the postulation of an unlimited range of new kinds of meson fields each having almost any conceivable rest mass, charge, spin, magnetic moment, etc. And if such postulates should prove to be inadequate, it is conceivable that we may have to introduce nonlocal operators, nonlinear fields, S-matrices, etc. This means that when the theory is found to be inadequate (as now happens, for example, at distances of the order of 10^{-13} cm), it is always possible, and, in fact, usually quite natural, to assume that the theory can be made to agree with experiment by some as yet unknown change in the mathematical formulation alone, not requiring any fundamental changes in the physical interpretation. This means that as long as we accept the usual physical interpretation of the quantum theory, we cannot be led by any conceivable experiment to give up this interpretation, even if it should happen to be wrong. The usual physical interpretation therefore presents us with a considerable danger of falling into a trap, consisting of a self-closing chain of circular hypotheses, which are in principle unverifiable if true. The only way of avoiding the possibility of such a trap is to study the consequences of postulates that contradict assumptions (1) and (2) at the outset. Thus, we could, for example, postulate that the precise outcome of each individual measurement process is in principle determined by some at present "hidden" elements or variables; and we could then try to find experiments that depended in a unique and reproducible way on the assumed state of these hidden elements or variables. If such predictions are verified, we should then obtain experimental evidence favoring the hypothesis that hidden variables exist. If they are not verified, however, the correctness of the usual interpretation of the quantum theory is not necessarily proved, since it may be necessary instead to alter the specific character of the theory that is supposed to describe the behavior of the assumed hidden variables.

We conclude then that a choice of the present interpretation of the quantum theory involves a real physical limitation on the kinds of theories that we wish to take into consideration. From the arguments given here, however, it would seem that there are no secure experimental or theoretical grounds on which we can base such a choice because this choice follows from hypotheses that cannot conceivably be subjected to an experimental test and because we now have an alternative interpretation.

4. NEW PHYSICAL INTERPRETATION OF SCHROEDINGER'S EQUATION

We shall now give a general description of our suggested physical interpretation of the present mathematical formulation of the quantum theory. We shall carry out a more detailed description in subsequent sections of this paper.

We begin with the one-particle Schroedinger equation, and shall later generalize to an arbitrary number of particles. This wave equation is

$$i\hbar\partial\psi/\partial t=-(\hbar^2/2m)\nabla^2\psi+V(\mathbf{x})\psi. \tag{1}$$

Now ψ is a complex function, which can be expressed as

$$\psi=R\exp(iS/\hbar), \tag{2}$$

where R and S are real. We readily verify that the equations for R and S are

$$\frac{\partial R}{\partial t}=-\frac{1}{2m}[R\nabla^2 S+2\nabla R\cdot\nabla S], \tag{3}$$

$$\frac{\partial S}{\partial t}=-\left[\frac{(\nabla S)^2}{2m}+V(\mathbf{x})-\frac{\hbar^2}{2m}\frac{\nabla^2 R}{R}\right]. \tag{4}$$

It is convenient to write $P(\mathbf{x})=R^2(\mathbf{x})$, or $R=P^{\frac{1}{2}}$ where $P(\mathbf{x})$ is the probability density. We then obtain

$$\frac{\partial P}{\partial t}+\nabla\cdot\left(P\frac{\nabla S}{m}\right)=0, \tag{5}$$

$$\frac{\partial S}{\partial t}+\frac{(\nabla S)^2}{2m}+V(\mathbf{x})-\frac{\hbar^2}{4m}\left[\frac{\nabla^2 P}{P}-\frac{1}{2}\frac{(\nabla P)^2}{P^2}\right]=0. \tag{6}$$

Now, in the classical limit ($\hbar\to 0$) the above equations are subject to a very simple interpretation. The function $S(\mathbf{x})$ is a solution of the Hamilton-Jacobi equation. If we consider an ensemble of particle trajectories which are solutions of the equations of motion, then it is a well-known theorem of mechanics that if all of these trajectories are normal to any given surface of constant S, then they are normal to all surfaces of constant S, and $\nabla S(\mathbf{x})/m$ will be equal to the velocity vector, $\mathbf{v}(\mathbf{x})$, for any particle passing the point $\mathbf{x}$. Equation (5) can therefore be re-expressed as

$$\partial P/\partial t+\nabla\cdot(P\mathbf{v})=0. \tag{7}$$

This equation indicates that it is consistent to regard $P(\mathbf{x})$ as the probability density for particles in our ensemble. For in that case, we can regard $P\mathbf{v}$ as the mean current of particles in this ensemble, and Eq. (7) then simply expresses the conservation of probability.

Let us now see to what extent this interpretation can be given a meaning even when $\hbar\neq 0$. To do this, let us assume that each particle is acted on, not only by a "classical" potential, $V(\mathbf{x})$ but also by a "quantum-mechanical" potential,

$$U(\mathbf{x})=\frac{-\hbar^2}{4m}\left[\frac{\nabla^2 P}{P}-\frac{1}{2}\frac{(\nabla P)^2}{P^2}\right]=\frac{-\hbar^2}{2m}\frac{\nabla^2 R}{R}. \tag{8}$$

Then Eq. (6) can still be regarded as the Hamilton-Jacobi equation for our ensemble of particles, $\nabla S(\mathbf{x})/m$ can still be regarded as the particle velocity, and Eq. (5) can still be regarded as describing conservation of probability in our ensemble. Thus, it would seem that we have here the nucleus of an alternative interpretation for Schroedinger's equation.

The first step in developing this interpretation in a more explicit way is to associate with each electron a particle having precisely definable and continuously varying values of position and momentum. The solution of the modified Hamilton-Jacobi equation (4) defines an ensemble of possible trajectories for this particle, which can be obtained from the Hamilton-Jacobi function, $S(\mathbf{x})$, by integrating the velocity, $\mathbf{v}(\mathbf{x})=\nabla S(\mathbf{x})/m$. The equation for S implies, however, that the particles moves under the action of a force which is not entirely derivable from the classical potential, $V(\mathbf{x})$, but which also obtains a contribution from the "quantum-mechanical" potential, $U(\mathbf{x})=(-\hbar^2/2m)\times\nabla^2 R/R$. The function, $R(\mathbf{x})$, is not completely arbitrary, but is partially determined in terms of $S(\mathbf{x})$ by the differential Eq. (3). Thus R and S can be said to codetermine each other. The most convenient way of obtaining R and S is, in fact, usually to solve Eq. (1) for the Schroedinger wave function, ψ, and then to use the relations,

$$\psi=U+iW=R[\cos(S/\hbar)+i\sin(S/\hbar)],$$
$$R^2=U^2+V^2;\quad S=\hbar\tan^{-1}(W/U).$$

Since the force on a particle now depends on a function of the absolute value, $R(\mathbf{x})$, of the wave function, $\psi(\mathbf{x})$, evaluated at the actual location of the particle, we have effectively been led to regard the wave function of an individual electron as a mathematical representation of an objectively real field. This field exerts a force on the particle in a way that is analogous to, but not identical with, the way in which an electromagnetic field exerts a force on a charge, and a meson field exerts a force on a nucleon. In the last analysis, there is, of course, no reason why a particle should not be acted on by a ψ-field, as well as by an electromagnetic field, a gravitational field, a set of meson fields, and perhaps by still other fields that have not yet been discovered.

The analogy with the electromagnetic (and other) field goes quite far. For just as the electromagnetic field obeys Maxwell's equations, the ψ-field obeys Schroedinger's equation. In both cases, a complete specification of the fields at a given instant over every point in space determines the values of the fields for all times. In both cases, once we know the field functions, we can calculate force on a particle, so that, if we also know the initial position and momentum of the particle, we can calculate its entire trajectory.

In this connection, it is worth while to recall that the use of the Hamilton-Jacobi equation in solving for the motion of a particle is only a matter of convenience and that, in principle, we can always solve directly by using Newton's laws of motion and the correct boundary conditions. The equation of motion of a particle acted on by the classical potential, $V(\mathbf{x})$, and the "quantum-mechanical" potential, Eq. (8), is

$$md^2\mathbf{x}/dt^2=-\nabla\{V(\mathbf{x})-(\hbar^2/2m)\nabla^2 R/R\}. \tag{8a}$$

It is in connection with the boundary conditions appearing in the equations of motion that we find the only fundamental difference between the ψ-field and other fields, such as the electromagnetic field. For in order to obtain results that are equivalent to those of the usual interpretation of the quantum theory, we are required to restrict the value of the initial particle momentum to $\mathbf{p}=\nabla S(\mathbf{x})$. From the application of Hamilton-Jacobi theory to Eq. (6), it follows that this restriction is consistent, in the sense that if it holds initially, it will hold for all time. Our suggested new interpretation of the quantum theory implies, however, that this restriction is not inherent in the conceptual structure. We shall see in Sec. 9, for example, that it is

quite consistent in our interpretation to contemplate modifications in the theory, which permit an arbitrary relation between $\mathbf{p}$ and $\nabla S(\mathbf{x})$. The law of force on the particle can, however, be so chosen that in the atomic domain, $\mathbf{p}$ turns out to be very nearly equal to $\nabla S(\mathbf{x})/m$, while in processes involving very small distances, these two quantities may be very different. In this way, we can improve the analogy between the ψ-field and the electromagnetic field (as well as between quantum mechanics and classical mechanics).

Another important difference between the ψ-field and the electromagnetic field is that, whereas Schroedinger's equation is homogeneous in ψ, Maxwell's equations are inhomogeneous in the electric and magnetic fields. Since inhomogeneities are needed to give rise to radiation, this means that our present equations imply that the ψ-field is not radiated or absorbed, but simply changes its form while its integrated intensity remains constant. This restriction to a homogeneous equation is, however, like the restriction to a homogeneous equation is, however, like the restriction to $\mathbf{p}=\nabla S(\mathbf{x})$, not inherent in the conceptual structure of our new interpretation. Thus, in Sec. 9, we shall show that one can consistently postulate inhomogeneities in the equation governing ψ, which produce important effects only at very small distances, and negligible effects in the atomic domain. If such inhomogeneities are actually present, then the ψ-field will be subject to being emitted and absorbed, but only in connection with processes associated with very small distances. Once the ψ-field has been emitted, however, it will in all atomic processes simply obey Schroedinger's equation as a very good approximation. Nevertheless, at very small distances, the value of the ψ-field would, as in the case of the electromagnetic field, depend to some extent on the actual location of the particle.

Let us now consider the meaning of the assumption of a statistical ensemble of particles with a probability density equal to $P(\mathbf{x})=R^2(\mathbf{x})=|\psi(\mathbf{x})|^2$. From Eq. (5), it follows that this assumption is consistent, provided that ψ satisfies Schroedinger's equation, and $\mathbf{v}=\nabla S(\mathbf{x})/m$. This probability density is numerically equal to the probability density of particles obtained in the usual interpretation. In the usual interpretation, however, the need for a probability description is regarded as inherent in the very structure of matter (see Sec. 2), whereas in our interpretation, it arises, as we shall see in Paper II, because from one measurement to the next, we cannot in practice predict or control the precise location of a particle, as a result of corresponding unpredictable and uncontrollable disturbances introduced by the measuring apparatus. Thus, in our interpretation, the use of a statistical ensemble is (as in the case of classical statistical mechanics) only a practical necessity, and not a reflection of an inherent limitation on the precision with which it is correct for us to conceive of the variables defining the state of the system. Moreover, it is clear that if in connection with very small distances we are ultimately required to give up the special assumptions that ψ satisfies Schroedinger's equation and that $\mathbf{v}=\nabla S(\mathbf{x})/m$, then $|\psi|^2$ will cease to satisfy a conservation equation and will therefore also cease to be able to represent the probability density of particles. Nevertheless, there would still be a true probability density of particles which is conserved. Thus, it would become possible in principle to find experiments in which $|\psi|^2$ could be distinguished from the probability density, and therefore to prove that the usual interpretation, which gives $|\psi|^2$ only a probability interpretation must be inadequate. Moreover, we shall see in Paper II that with the aid of such modifications in the theory, we could in principle measure the particle positions and momenta precisely, and thus violate the uncertainty principle. As long as we restrict ourselves to conditions in which Schroedinger's equation is satisfied, and in which $\mathbf{v}=\nabla S(\mathbf{x})/m$, however, the uncertainty principle will remain an effective practical limitation on the possible precision of measurements. This means that at present, the particle positions and momenta should be regarded as "hidden" variables, since as we shall see in Paper II, we are not now able to obtain experiments that localize them to a region smaller than that in which the intensity of the ψ-field is appreciable. Thus, we cannot yet find clear-cut experimental proof that the assumption of these variables is necessary, although it is entirely possible that, in the domain of very small distances, new modifications in the theory may have to be introduced, which would permit a proof of the existence of the definite particle position and momentum to be obtained.

We conclude that our suggested interpretation of the quantum theory provides a much broader conceptual framework than that provided by the usual interpretation, for all of the results of the usual interpretation are obtained from our interpretation if we make the following three special assumptions which are mutually consistent:

(1) That the ψ-field satisfies Schroedinger's equation.
(2) That the particle momentum is restricted to $\mathbf{p}=\nabla S(\mathbf{x})$.
(3) That we do not predict or control the precise location of the particle, but have, in practice, a statistical ensemble with probability density $P(\mathbf{x})=|\psi(\mathbf{x})|^2$. The use of statistics is, however, not inherent in the conceptual structure, but merely a consequence of our ignorance of the precise initial conditions of the particle.

As we shall see in Sec. 9, it is entirely possible that a better theory of phenomena involving distances of the order of 10^{-13} cm or less would require us to go beyond the limitations of these special assumptions. Our principal purpose in this paper (and in Paper II) is to show, however, that if one makes these special assumptions, our interpretation leads in all possible experiments to the same predictions as are obtained from the usual interpretation.[9]

It is now easy to understand why the adoption of the

usual interpretation of the quantum theory would tend to lead us away from the direction of our suggested alternative interpretation. For in a theory involving hidden variables, one would normally expect that the behavior of an individual system should not depend on the statistical ensemble of which it is a member, because this ensemble refers to a series of similar but disconnected experiments carried out under equivalent initial conditions. In our interpretation, however, the "quantum-mechanical" potential, $U(\mathbf{x})$, acting on an individual particle depends on a wave intensity, $P(\mathbf{x})$, that is also numerically equal to a probability density in our ensemble. In the terminology of the usual interpretation of the quantum theory, in which one tacitly assumes that the wave function has only one interpretation; namely, in terms of a probability, our suggested new interpretation would look like a mysterious dependence of the individual on the statistical ensemble of which it is a member. In our interpretation, such a dependence is perfectly rational, because the wave function can consistently be interpreted both as a force and as a probability density.[12]

It is instructive to carry our analogy between the Schroedinger field and other kinds of fields a bit further. To do this, we can derive the wave Eqs. (5) and (6) from a Hamiltonian functional. We begin by writing down the expression for the mean energy as it is expressed in the usual quantum theory:

$$\bar{H}=\int\psi^*\left(-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{x})\right)\psi d\mathbf{x}$$

$$=\int\left\{\frac{\hbar^2}{2m}|\nabla\psi|^2+V(\mathbf{x})|\psi|^2\right\}d\mathbf{x}.$$

Writing $\psi=P^{\frac{1}{2}}\exp(iS/\hbar)$, we obtain

$$\bar{H}=\int P(\mathbf{x})\left\{\frac{(\nabla S)^2}{2m}+V(\mathbf{x})+\frac{\hbar^2}{8m}\frac{(\nabla P)^2}{P^2}\right\}d\mathbf{x}. \quad (9)$$

We shall now reinterpret $P(\mathbf{x})$ as a field coordinate, defined at each point, $\mathbf{x}$, and we shall tentatively assume that $S(\mathbf{x})$ is the momentum, canonically conjugate to $P(\mathbf{x})$. That such an assumption is appropriate can be verified by finding the Hamiltonian equations of motion for $P(\mathbf{x})$ and $S(\mathbf{x})$, under the assumption that the Hamiltonian functional is equal to $\bar{H}$ (See Eq. (9)). These equations of motion are

$$\dot{P}=\frac{\delta\bar{H}}{\delta S}=-\frac{1}{m}\nabla\cdot(P\nabla S),$$

$$\dot{S}=-\frac{\delta\bar{H}}{\delta P}=-\left[\frac{(\nabla S)^2}{2m}+V(\mathbf{x})-\frac{\hbar^2}{4m}\left(\frac{\nabla^2P}{P}-\frac{1}{2}\frac{(\nabla P)^2}{P^2}\right)\right].$$

[12] This consistency is guaranteed by the conservation Eq. (7). The questions of why an arbitrary statistical ensemble tends to decay into an ensemble with a probability density equal to $\psi^*\psi$ will be discussed in Paper II, Sec. 7.

These are, however, the same as the correct wave Eqs. (5) and (6).

We can now show that the mean particle energy averaged over our ensemble is equal to the usual quantum mechanical mean value of the Hamiltonian, $\bar{H}$. To do this, we note that according to Eqs. (3) and (6), the energy of a particle is

$$E(\mathbf{x})=-\frac{\partial S(\mathbf{x})}{\partial t}=\left[\frac{(\nabla S)^2}{2m}+V(\mathbf{x})-\frac{\hbar^2}{2m}\frac{\nabla^2R}{R}\right]. \quad (10)$$

The mean particle energy is found by averaging $E(\mathbf{x})$ with the weighting function, $P(\mathbf{x})$. We obtain

$$\langle E\rangle_{\substack{\text{ensemble}\\\text{average}}}=\int P(\mathbf{x})E(\mathbf{x})d\mathbf{x}$$

$$=\int P(\mathbf{x})\left[\frac{(\nabla S)^2}{2m}+V(\mathbf{x})\right]d\mathbf{x}-\frac{\hbar^2}{2m}\int R\nabla^2Rd\mathbf{x}.$$

A little integration by parts yields

$$\langle E\rangle_{\substack{\text{ensemble}\\\text{average}}}=\int P(\mathbf{x})\left[\frac{(\nabla S)^2}{2m}+V(\mathbf{x})+\frac{\hbar^2}{8m}\frac{(\nabla P)^2}{P^2}\right]d\mathbf{x}=\bar{H}. \quad (11)$$

5. THE STATIONARY STATE

We shall now show how the problem of stationary states is to be treated in our interpretation of the quantum theory.

The following seem to be reasonable requirements in our interpretation for a stationary state:

(1) The particle energy should be a constant of the motion.

(2) The quantum-mechanical potential should be independent of time.

(3) The probability density in our statistical ensemble should be independent of time.

It is easily verified that these requirements can be satisfied with the assumption that

$$\psi(\mathbf{x},t)=\psi_0(\mathbf{x})\exp(-iEt/\hbar)$$

$$=R_0(\mathbf{x})\exp[i(\Phi(\mathbf{x})-Et)/\hbar]. \quad (12)$$

From the above, we obtain $S=\Phi(\mathbf{x})-Et$. According to the generalized Hamilton-Jacobi Eq. (4), the particle energy is given by

$$\partial S/\partial t=-E.$$

Thus, we verify that the particle energy is a constant of the motion. Moreover, since $P=R^2=|\psi|^2$, it follows that P (and R) are independent of time. This means that both the probability density in our ensemble and the quantum-mechanical potential are also time independent.

The reader will readily verify that no other form of solution of Schroedinger's equation will satisfy all three of our criteria for a stationary state.

Since ψ is now being regarded as a mathematical representation of an objectively real force field, it follows that (like the electromagnetic field) it should be everywhere finite, continuous, and single valued. These requirements will guarantee in all cases that occur in practice that the allowed values of the energy in a stationary state, and the corresponding eigenfunctions are the same as are obtained from the usual interpretation of the theory.

In order to show in more detail what a stationary state means in our interpretation, we shall now consider three examples of stationary states.

Case 1: "s" State

The first case that we shall consider is an "s" state. In an "s" state, the wave function is

$$\psi=f(r)\exp[i(\alpha-Et)/\hbar], \tag{13}$$

where α is an arbitrary constant and r is the radius taken from the center of the atom. We conclude that the Hamilton-Jacobi function is

$$S=\alpha-Et.$$

The particle velocity is

$$\mathbf{v}=\nabla S=0.$$

The particle is therefore simply standing still, wherever it may happen to be. How can it do this? The absence of motion is possible because the applied force, $-\nabla V(\mathbf{x})$, is balanced by the "quantum-mechanical" force, $(\hbar^2/2m)\nabla(\nabla^2R/R)$, produced by the Schroedinger ψ-field acting on its own particle. There is, however, a statistical ensemble of possible positions of the particle, with a probability density, $P(\mathbf{x})=(f(r))^2$.

Case 2: State with Nonzero Angular Momentum

In a typical state of nonzero angular momentum, we have

$$\psi=f_n{}^l(r)P_l{}^m(\cos\theta)\exp[i(\beta-Et+\hbar m\phi)/\hbar], \tag{14}$$

where θ and ϕ are the colatitude and azimuthal polar angles, respectively, $P_l{}^m$ is the associated Legendre polynomial, and β is a constant. The Hamilton-Jacobi function is $S=\beta-Et+\hbar m\phi$. From this result it follows that the z component of the angular momentum is equal to $\hbar m$. To prove this, we write

$$L_z=xp_y-yp_x=x\partial S/\partial y-y\partial S/\partial x=\partial S/\partial\phi=\hbar m. \tag{15}$$

Thus, we obtain a statistical ensemble of trajectories which can have different forms, but all have the same "quantized" value of the z component of the angular momentum.

Case 3: A Scattering Problem

Let us now consider a scattering problem. Because it is comparatively easy to analyze, we shall discuss a hypothetical experiment, in which an electron is incident in the z direction with an initial momentum, p_0, on a system consisting of two slits.[13] After the electron passes through the slit system, its position is measured and recorded, for example, on a photographic plate.

Now, in the usual interpretation of the quantum theory, the electron is described by a wave function. The incident part of the wave function is $\psi_0\sim\exp(ip_0z/\hbar)$; but when the wave passes through the slit system, it is modified by interference and diffraction effects, so that it will develop a characteristic intensity pattern by the time it reaches the position measuring instrument. The probability that the electron will be detected between $\mathbf{x}$ and $\mathbf{x}+d\mathbf{x}$ is $|\psi(\mathbf{x})|^2d\mathbf{x}$. If the experiment is repeated many times under equivalent initial conditions, one eventually obtains a pattern of hits on the photographic plate that is very reminiscent of the interference patterns of optics.

In the usual interpretation of the quantum theory, the origin of this interference pattern is very difficult to understand. For there may be certain points where the wave function is zero when both slits are open, but not zero when only one slit is open. How can the opening of a second slit prevent the electron from reaching certain points that it could reach if this slit were closed? If the electron acted completely like a classical particle, this phenomenon could not be explained at all. Clearly, then the wave aspects of the electron must have something to do with the production of the interference pattern. Yet, the electron cannot be identical with its associated wave, because the latter spreads out over a wide region. On the other hand, when the electron's position is measured, it always appears at the detector as if it were a localized particle.

The usual interpretation of the quantum theory not only makes no attempt to provide a single precisely defined conceptual model for the production of the phenomena described above, but it asserts that no such model is even conceivable. Instead of a single precisely defined conceptual model, it provides, as pointed out in Sec. 2, a pair of complementary models, *viz.*, particle and wave, each of which can be made more precise only under conditions which necessitate a reciprocal decrease in the degree of precision of the other. Thus, while the electron goes through the slit system, its position is said to be inherently ambiguous, so that if we wish to obtain an interference pattern, it is meaningless to ask through which slit an individual electron actually passed. Within the domain of space within which the position of the electron has no meaning we can use the wave model and thus describe the subsequent production of interference. If, however, we

[13] This experiment is discussed in some detail in reference 2, Chapter 6, Sec. 2.

tried to define the position of the electron as it passed the slit system more accurately by means of a measurement, the resulting disturbance of its motion produced by the measuring apparatus would destroy the interference pattern. Thus, conditions would be created in which the particle model becomes more precisely defined at the expense of a corresponding decrease in the degree of definition of the wave model. When the position of the electron is measured at the photographic plate, a similar sharpening of the degree of definition of the particle model occurs at the expense of that of the wave model.

In our interpretation of the quantum theory, this experiment is described causally and continuously in terms of a single precisely definable conceptual model. As we have already shown, we must use the same wave function as is used in the usual interpretation; but instead we regard it as a mathematical representation of an objectively real field that determines part of the force acting on the particle. The initial momentum of the particle is obtained from the incident wave function, $\exp(ip_0z/\hbar)$, as $p=\partial s/\partial z=p_0$. We do not in practice, however, control the initial location of the particle, so that although it goes through a definite slit, we cannot predict which slit this will be. The particle is at all times acted on by the "quantum-mechanical" potential, $U=(-\hbar^2/2m)\nabla^2R/R$. While the particle is incident, this potential vanishes because R is then a constant; but after it passes through the slit system, the particle encounters a quantum-mechanical potential that changes rapidly with position. The subsequent motion of the particle may therefore become quite complicated. Nevertheless, the probability that a particle shall enter a given region, $d\mathbf{x}$, is as in the usual interpretation, equal to $|\psi(\mathbf{x})|^2d\mathbf{x}$. We therefore deduce that the particle can never reach a point where the wave function vanishes. The reason is that the "quantum-mechanical" potential, U, becomes infinite when R becomes zero. If the approach to infinity happens to be through positive values of U, there will be an infinite force repelling the particle away from the origin. If the approach is through negative values of U, the particle will go through this point with infinite speed, and thus spend no time there. In either case, we obtain a simple and precisely definable conceptual model explaining why particles can never be found at points where the wave function vanishes.

If one of the slits is closed, the "quantum-mechanical" potential is correspondingly altered, because the ψ-field is changed, and the particle may then be able to reach certain points which it was unable to reach when both slits were open. The slit is therefore able to affect the motion of the particle only indirectly, through its effect on the Schroedinger ψ-field. Moreover, as we shall see in Paper II, if the position of the electron is measured while it is passing through the slit system, the measuring apparatus will, as in the usual interpretation, create a disturbance that destroys the interference pattern. In our interpretation, however, the necessity for this destruction is not inherent in the conceptual structure; and as we shall see, the destruction of the interference pattern could in principle be avoided by means of other ways of making measurements, ways which are conceivable but not now actually possible.

6. THE MANY-BODY PROBLEM

We shall now extend our interpretation of the quantum theory to the problem of many bodies. We begin with the Schroedinger equation for two particles. (For simplicity, we assume that they have equal masses, but the extension of our treatment to arbitrary masses will be obvious.)

$$i\hbar\frac{\partial\psi}{\partial t}=-\frac{\hbar^2}{2m}(\nabla_1^2\psi+\nabla_2^2\psi)+V(\mathbf{x}_1,\mathbf{x}_2)\psi.$$

Writing $\psi=R(\mathbf{x}_1,\mathbf{x}_2)\exp[iS(\mathbf{x}_1,\mathbf{x}_2)/\hbar]$ and $R^2=P$, we obtain

$$\frac{\partial P}{\partial t}+\frac{1}{m}[\nabla_1\cdot P\nabla_1S+\nabla_2\cdot P\nabla_2S]=0, \tag{16}$$

$$\frac{\partial S}{\partial t}+\frac{(\nabla_1S)^2+(\nabla_2S)^2}{2m}+V(\mathbf{x}_1,\mathbf{x}_2)-\frac{\hbar^2}{2mR}[\nabla_1^2R+\nabla_2^2R]=0. \tag{17}$$

The above equations are simply a six-dimensional generalization of the similar three-dimensional Eqs. (5) and (6) associated with the one-body problem. In the two-body problem, the system is described therefore by a six-dimensional Schroedinger wave and by a six-dimensional trajectory, specifying the actual location of each of the two particles. The velocity of this trajectory has components, ∇_1S/m and ∇_2S/m, respectively, in each of the three-dimensional surfaces associated with a given particle. $P(\mathbf{x}_1,\mathbf{x}_2)$ then has a dual interpretation. First, it defines a "quantum-mechanical" potential, acting on each particle

$$U(\mathbf{x}_1,\mathbf{x}_2)=-(\hbar^2/2mR)[\nabla_1^2R+\nabla_2^2R].$$

This potential introduces an additional effective interaction between particles over and above that due to the classically inferrable potential $V(\mathbf{x})$. Secondly, the function $P(\mathbf{x}_1,\mathbf{x}_2)$ can consistently be regarded as the probability density of representative points $(\mathbf{x}_1,\mathbf{x}_2)$ in our six-dimensional ensemble.

The extension to an arbitrary number of particles is straightforward, and we shall quote only the results here. We introduce the wave function, $\psi=R(\mathbf{x}_1,\mathbf{x}_2,\cdots\mathbf{x}_n)\exp[iS(\mathbf{x}_1,\mathbf{x}_2\cdots\mathbf{x}_n)/\hbar]$ and define a $3n$-dimensional trajectory, where n is the number of particles, which describes the behavior of every particle in the system. The velocity of the ith particle is $\mathbf{v}_i=\nabla_iS(\mathbf{x}_1,\mathbf{x}_2\cdots\mathbf{x}_n)/m$. The function $P(\mathbf{x}_1,\mathbf{x}_2\cdots\mathbf{x}_n)=R^2$ has two

interpretations. First, it defines a "quantum-mechanical" potential

$$U(\mathbf{x}_1, \mathbf{x}_2 \cdots \mathbf{x}_n) = -\frac{\hbar^2}{2mR} \sum_{i=1}^{n} \nabla_i^2 R(\mathbf{x}_1, \mathbf{x}_2 \cdots \mathbf{x}_n). \quad (18)$$

Secondly, $P(\mathbf{x}_1, \mathbf{x}_2 \cdots \mathbf{x}_n)$ is equal to the density of representative points $(\mathbf{x}_1, \mathbf{x}_2 \cdots \mathbf{x}_n)$ in our $3n$-dimensional ensemble.

We see here that the "effective potential," $U(\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_n)$, acting on a particle is equivalent to that produced by a "many-body" force, since the force between any two particles may depend significantly on the location of every other particle in the system. An example of the effects of such a force is given by the exclusion principle. Thus, if the wave function is antisymmetric, we deduce that the "quantum-mechanical" forces will be such as to prevent two particles from ever reaching the same point in space, for in this case, we must have $P=0$.

7. TRANSITIONS BETWEEN STATIONARY STATES—THE FRANCK-HERTZ EXPERIMENT

Our interpretation of the quantum theory describes all processes as basically causal and continuous. How then can it lead to a correct description of processes such as the Franck-Hertz experiment, the photoelectric effect, and the Compton effect, which seem to call most strikingly for an interpretation in terms of discontinuous and incompletely determined transfers of energy and momentum? In this section, we shall answer this question by applying our suggested interpretation of the quantum theory in the analysis of the Franck-Hertz experiment. Here, we shall see that the apparently discontinuous nature of the process of transfer of energy from the bombarding particle to the atomic electron is brought about by the "quantum-mechanical" potential, $U=(-\hbar^2/2m)\nabla^2 R/R$, which does not necessarily become small when the wave intensity becomes small. Thus, even if the force of interaction between two particles is very weak, so that a correspondingly small disturbance of the Schroedinger wave function is produced by the interaction of these particles, this disturbance is capable of bringing about very large transfers of energy and momentum between the particles in a very short time. This means that if we view only the end results, this process presents the aspect of being discontinuous. Moveover, we shall see that the precise value of the energy transfer is in principle determined by the initial position of each particle and by the initial form of the wave function. Since we cannot in practice predict or control the initial particle positions with complete precision, we are also unable to predict or control the final outcome of such an experiment, and can, in practice, predict only the probability of a given outcome. Because the probability that the particles will enter a region with coordinates, $\mathbf{x}_1$, $\mathbf{x}_2$, is proportional to $R^2(\mathbf{x}_1, \mathbf{x}_2)$, we conclude that although a Schroedinger wave of low intensity can bring about large transfers of energy, such a process is (as in the usual interpretation) highly improbable.

In Appendix A of Paper II, we shall see that similar possibilities arise in connection with the interaction of the electromagnetic field with charged matter, so that electromagnetic waves can very rapidly transfer a full quantum of energy (and momentum) to an electron, even after they have spread out and fallen to a very low intensity. In this way, we shall explain the photoelectric effect and the Compton effect. Thus, we are able in our interpretation to understand by means of a causal and continuous model just those properties of matter and light which seem most convincingly to require the assumption of discontinuity and incomplete determinism.

Before we discuss the process of interaction between two particles, we shall find it convenient to analyze the problem of an isolated single particle that happens to be in a nonstationary state. Because the field function ψ is a solution of Schroedinger's equation, we can linearly suppose stationary-state solutions of this equation and in this way obtain new solutions. As an illustration, let us consider a superposition of two solutions

$$\psi = C_1\psi_1(\mathbf{x})\exp(-iE_1t/\hbar) + C_2\psi_2(\mathbf{x})\exp(-iE_2t/\hbar),$$

where C_1, C_2, ψ_1, and ψ_2 are real. Thus we write $\psi_1 = R_1$, $\psi_2 = R_2$, and

$$\psi = \exp[-i(E_1+E_2)t/2\hbar]\{C_1R_1\exp[-i(E_1-E_2)t/2\hbar] + C_2R_2\exp[i(E_1-E_2)t/2\hbar]\}.$$

Writing $\psi = R\exp(iS/\hbar)$, we obtain

$$R^2 = C_1^2R_1^2(\mathbf{x}) + C_2^2R_2^2(\mathbf{x}) + 2C_1C_2R_1(\mathbf{x})R_2(\mathbf{x})\cos[(E_1-E_2)t/2\hbar], \quad (19)$$

$$\tan\left\{\frac{S+(E_1-E_2)t/2}{\hbar}\right\} = \frac{C_2R_2(\mathbf{x})-C_1R_1(\mathbf{x})}{C_2R_2(\mathbf{x})+C_1R_1(\mathbf{x})}\tan\left\{\frac{(E_1-E_2)t}{2\hbar}\right\}. \quad (20)$$

We see immediately that the particle experiences a "quantum-mechanical" potential, $U(\mathbf{x})=(-\hbar/2m)\nabla^2R/R$, which fluctuates with angular frequency, $w=(E_1-E_2)/\hbar$, and that the energy of this particle, $E=-\partial S/\partial t$, and its momentum $\mathbf{p}=\nabla S$, fluctuate with the same angular frequency. If the particle happens to enter a region of space where R is small, these fluctuations can become quite violent. We see then that, in general, the orbit of a particle in a nonstationary state is very irregular and complicated, resembling Brownian motion more closely than it resembles the smooth track of a planet around the sun.

If the system is isolated, these fluctuations will continue forever. The result is quite reasonable, since as is well known, a system can make a transition from one stationary state to another only if it can exchange en-

ergy with some other system. In order to treat the problem of transition between stationary states, we must therefore introduce another system capable of exchanging energy with the system of interest. In this section, we shall discuss the Franck-Hertz experiment, in which this other system consists of a bombarding particle. For the sake of illustration, let us suppose that we have hydrogen atoms of energy E_0 and wave function, $\psi_0(\mathbf{x})$, which are bombarded by particles that can be scattered inelastically, leaving the atom with energy E_n and wave function, $\psi_n(\mathbf{x})$.

We begin by writing down the initial wave function, $\Psi_i(\mathbf{x}, \mathbf{y}, t)$. The incident particle, whose coordinates are represented by $\mathbf{y}$ must be associated with a wave packet, which can be written as

$$f_0(\mathbf{y}, t)=\int e^{i\mathbf{k}\cdot\mathbf{y}} f(\mathbf{k}-\mathbf{k}_0)\exp(-i\hbar k^2 t/2m)d\mathbf{k}. \quad (21)$$

The center of this packet occurs where the phase has an extremum as a function of $\mathbf{k}$, or where $\mathbf{y}=\hbar\mathbf{k}_0 t/m$.

Now, as in the usual interpretation, we begin by writing the incident wave function for the combined system as a product

$$\Psi_i=\psi_0(\mathbf{x})\exp(-iE_0 t/\hbar)f_0(\mathbf{y}, t). \quad (22)$$

Let us now see how this wave function is to be understood in our interpretation of the theory. As pointed out in Sec. 6, the wave function is to be regarded as a mathematical representation of a six-dimensional but objectively real field, capable of producing forces that act on the particles. We also assume a six-dimensional representative point, described by the coordinates of the two particles, $\mathbf{x}$ and $\mathbf{y}$. We shall now see that when the combined wave function takes the form (22) involving a product of a function of $\mathbf{x}$ and a function of $\mathbf{y}$, the six-dimensional system can correctly be regarded as being made up of two independent three-dimentional subsystems. To prove this, we write

$$\psi_0(\mathbf{x})=R_0(\mathbf{x})\exp[iS_0(\mathbf{x})/\hbar] \quad \text{and}$$
$$f_0(\mathbf{y}, t)=M_0(\mathbf{y}, t)\exp[iN_0(\mathbf{y}, t)/\hbar].$$

We then obtain for the particle velocities

$$d\mathbf{x}/dt=(1/m)\nabla S_0(\mathbf{x}); \quad d\mathbf{y}/dt=(1/m)\nabla N_0(y, t), \quad (23)$$

and for the "quantum-mechanical" potential

$$U=-\frac{\hbar^2\{(\nabla_x^2+\nabla_y^2)R(\mathbf{x}, \mathbf{y})\}}{2mR(\mathbf{x}, \mathbf{y})}$$
$$=\frac{-\hbar^2}{2m}\left\{\frac{\nabla^2 R_0(\mathbf{x})}{R_0(\mathbf{x})}+\frac{\nabla^2 M_0(\mathbf{y}, t)}{M_0(\mathbf{y}, t)}\right\}. \quad (24)$$

Thus, the particle velocities are independent and the "quantum-mechanical" potential reduces to a sum of terms, one involving only $\mathbf{x}$ and the other involving only $\mathbf{y}$. This means that the particles move independently. Moreover, the probability density, $P=R_0^2(\mathbf{x})\times M_0^2(\mathbf{y}, t)$, is a product of a function of $\mathbf{x}$ and a function of $\mathbf{y}$, indicating that the distribution in $\mathbf{x}$ is statistically independent of that in $\mathbf{y}$. We conclude, then, that whenever the wave function can be expressed as a product of two factors, each involving only the coordinates of a single system, then the two systems are completely independent of each other.

As soon as the wave packet in $\mathbf{y}$ space reaches the neighborhood of the atom, the two systems begin to interact. If we solve Schroedinger's equation for the combined system, we obtain a wave function that can be expressed in terms of the following series:

$$\Psi=\Psi_i+\sum_n\psi_n(\mathbf{x})\exp(-iE_n t/\hbar)f_n(\mathbf{y}, t), \quad (25)$$

where the $f_n(\mathbf{y}, t)$ are the expansion coefficients of the complete set of functions, $\psi_n(\mathbf{x})$. The asymptotic form of the wave function is[14]

$$\Psi=\Psi_i(\mathbf{x}, \mathbf{y})+\sum_n\psi_n(\mathbf{x})\exp\left(-\frac{iE_n t}{\hbar}\right)\int f(\mathbf{k}-\mathbf{k}_0)$$
$$\times\frac{\exp[ik_n\cdot\mathbf{r}-(\hbar k_n^2/2n)t]}{r}g_n(\theta, \phi, \mathbf{k})d\mathbf{k}, \quad (26)$$

where

$$\hbar^2 k_n^2/2m=(\hbar^2 k_0^2/2m)+E_0-E_n$$
(conservation of energy). (27)

The additional terms in the above equation represent outgoing wave packets, in which the particle speed, $\hbar k_n/m$, is correlated with the wave function, $\psi_n(\mathbf{x})$, representing the state in which the hydrogen atom is left. The center of the nth packet occurs at

$$r_n=(\hbar k_n/m)t. \quad (28)$$

It is clear that because the speed depends on the hydrogen atom quantum number, n, every one of these packets will eventually be separated by distances which are so large that this separation is classically describable.

When the wave function takes the form (25), the two particles system must be described as a single six-dimensional system and not as a sum of two independent three-dimensional subsystems, for at this time, if we try to express the wave function as $\psi(\mathbf{x}, \mathbf{y})=R(\mathbf{x}, \mathbf{y})\times\exp[iS(\mathbf{x}, \mathbf{y})/\hbar]$, we find that the resulting expressions for R and S depend on $\mathbf{x}$ and $\mathbf{y}$ in a very complicated way. The particle momenta, $\mathbf{p}_1=\nabla_x S(\mathbf{x}, \mathbf{y})$ and $\mathbf{p}_2=\nabla_y S(\mathbf{x}, \mathbf{y})$, therefore become inextricably interdependent. The "quantum-mechanical" potential,

$$U=-\frac{\hbar^2}{2mR(\mathbf{x}, \mathbf{y})}(\nabla_x^2 R+\nabla_y^2 R)$$

ceases to be expressible as the sum of a term involving $\mathbf{x}$ and a term involving $\mathbf{y}$. The probability density,

[14] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions* (Clarendon Press, Oxford, 1933).

$R^2(\mathbf{x}, \mathbf{y})$ can no longer be written as a product of a function of $\mathbf{x}$ and a function of $\mathbf{y}$, from which we conclude that the probability distributions of the two particles are no longer statistically independent. Moreover, the motion of the particle is exceedingly complicated, because the expressions for R and S are somewhat analogous to those obtained in the simpler problem of a nonstationary state of a single particle [see Eqs. (19) and (20)]. In the region where the scattered waves $\psi_n(\mathbf{x})f_n(\mathbf{y}, t)$ have an amplitude comparable with that of the incident wave, $\psi_0(\mathbf{x})f_0(\mathbf{y}, t)$, the functions R and S, and therefore the "quantum-mechanical" potential and the particle momenta, undergo rapid and violent fluctuations, both as functions of position and of time. Because the quantum-mechanical potential has $R(\mathbf{x}, \mathbf{y}, t)$ in the denominator, these fluctuations may become very large in this region where R is small. If the particles happen to enter such a region, they may exchange very large quantities of energy and momentum in a very short time, even if the classical potential, $V(\mathbf{x}, \mathbf{y})$ is very small. A small value of $V(\mathbf{x}, \mathbf{y})$ implies, however, a correspondingly small value of the scattered wave amplitudes, $f_n(\mathbf{y}, t)$. Since the fluctuations become large only in the region where the scattered wave amplitude is comparable with the incident wave amplitude and since the probability that the particles shall enter a given region of $\mathbf{x}, \mathbf{y}$ space is proportional to $R^2(\mathbf{x}, \mathbf{y})$, it is clear that a large transfer of energy is improbable (although still always possible) when $V(\mathbf{x}, \mathbf{y})$ is small.

While interaction between the two particles takes place then, their orbits are subject to wild fluctuations. Eventually, however, the behavior of the system quiets down and becomes simple again. For after the wave function takes its asymptotic form (26), and the packets corresponding to different values of n have obtained classically describable separations, we can deduce that because the probability density is $|\psi|^2$, the outgoing particle must enter one of these packets and stay with that packet thereafter (since it does not enter the space between packets in which the probability density is negligibly different from zero). In the calculation of the particle velocities, $\mathbf{V}_1=\nabla_x S/m$, $\mathbf{V}_2=\nabla_y S/m$, and of the quantum-mechanical potential, $U=(-\hbar^2/2mR)(\nabla_x^2 R+\nabla_y^2 R)$, we can therefore ignore all parts of the wave function other than the one actually containing the outgoing particle. It follows that the system acts as if it had the wave function

$$\Psi_n=\psi_n(\mathbf{x})\exp\left(\frac{iE_n t}{\hbar}\right)\int f(\mathbf{k}-\mathbf{k}_0)\times\frac{\exp\{i[\mathbf{k}_n\cdot\mathbf{r}-(\hbar k_n^2 t/2m)t]\}}{r}g_n(\theta,\phi,\mathbf{k})d\mathbf{k}, \quad (29)$$

where n denotes the packet actually containing the outgoing particle. This means that for all practical purposes the complete wave function (26) of the system may be replaced by Eq. (29), which corresponds to an atomic electron in its nth quantum state, and to an outgoing particle with a correlated energy, $E_n'=\hbar^2k_n^2/2m$. Because the wave function is a product of a function of $\mathbf{x}$ and a function of $\mathbf{y}$, each system once again acts independently of the other. The wave function can now be renormalized because the multiplication of Ψ_n by a constant changes no physically significant quantity, such as the particle velocity or the "quantum-mechanical" potential. As shown in Sec. 5, when the electronic wave function is $\psi_n(\mathbf{x})\exp(-iE_n t/\hbar)$, its energy must be E_n. Thus, we have obtained a description of how it comes about that the energy is always transferred in quanta of size E_n-E_0.

It should be noted that while the wave packets are still separating, the electron energy is not quantized, but has a continuous range of values, which fluctuate rapidly. It is only the final value of the energy, appearing after the interaction is over that must be quantized. A similar result is obtained in the usual interpretation if one notes that because of the uncertainty principle, the energy of either system can become definite only after enough time has elapsed to complete the scattering process.[15]

In principle, the actual packet entered by the outgoing particle could be predicted if we knew the initial position of both particles and, of course, the initial form of the wave function of the combined system.[16] In practice, however, the particle orbits are very complicated and very sensitively dependent on the precise values of these initial positions. Since we do not at present know how to measure these initial positions precisely, we cannot actually predict the outcome of such an interaction process. The best that we can do is to predict the probability that an outgoing particle enters the nth packet within a given range of solid angle, $d\Omega$, leaving the hydrogen atom in its nth quantum state. In doing this, we use the fact that the probability density in $\mathbf{x}, \mathbf{y}$ space is $|\psi(\mathbf{x}, \mathbf{y})|^2$ and that as long as we are restricted to the nth packet, we can replace the complete wave function (26) by the wave function (29), corresponding to the packet that actually contains the particle. Now, by definition, we have $\int|\psi_n(\mathbf{x})|^2d\mathbf{x}=1$. The remaining integration of

$$\left|\int f(\mathbf{k}-\mathbf{k}_0)\frac{\exp\{i[k_n r-(\hbar k_n^2/2m)t]\}}{r}g_n(\theta,\phi,k)d\mathbf{k}\right|^2$$

over the region of space corresponding to the nth outgoing packet leads, however, to precisely the same probability of scattering as would have been obtained by applying the usual interpretation. We conclude, then, that if ψ satisfies Schroedinger's equation, that if $\mathbf{v}=\nabla S/m$, and that if the probability density of particles is $P(\mathbf{x}, \mathbf{y})=R^2(\mathbf{x}, \mathbf{y})$, we obtain in every respect

[15] See reference 2, Chapter 18, Sec. 19.

[16] Note that in the usual interpretation one assumes that *nothing* determines the precise outcome of an individual scattering process. Instead, one assumes that all descriptions are inherently and unavoidably statistical (see Sec. 2).

exactly the same physical predictions for this problem as are obtained when we use the usual interpretation.

There remains only one more problem; namely, to show that if the outgoing packets are subsequently brought together by some arrangement of matter that does not act on the atomic electron, the atomic electron and the scattered particle will continue to act independently.[17] To show that these two particles will continue to act independently, we note that in all practical applications, the outgoing particle soon interacts with some classically describable system. Such a system might consist, for example, of the host of atoms of the gas with which it collides or of the walls of a container. In any case, if the scattering process is ever to be observed, the outgoing particle must interact with a classically describable measuring apparatus. Now all classically describable systems have the property that they contain an enormous number of internal "thermodynamic" degrees of freedom that are inevitably excited when the outgoing particle interacts with the system. The wave function of the outgoing particle is then coupled to that of these internal thermodynamic degrees of freedom, which we represent as $y_1, y_2, \cdots y_s$. To denote this coupling, we write the wave function for the entire system as

$$\Psi=\sum_n \psi_n(\mathbf{x})\exp(-iE_nt/h)f_n(\mathbf{y}, y_1, y_2\cdots y_s). \quad (30)$$

Now, when the wave function takes this form, the overlapping of different packets in $\mathbf{y}$ space is not enough to produce interference between the different $\psi_n(\mathbf{x})$. To obtain such interference, it is necessary that the packets $f_n(\mathbf{y}, y_1, y_2, \cdots y_s)$ overlap in every one of the $S+3$ dimensions, $\mathbf{y}, y_1, y_2\cdots y_s$. The reader will readily convince himself, by considering a typical case such as a collision of the outgoing particle with a metal wall, that it is overwhelmingly improbable that two of the packets $f_n(\mathbf{y}_1, y_1, y_2\cdots y_s)$ will overlap with regard to every one of the internal thermodynamic coordinates, $y_1, y_2, \cdots y_s$, even if they are successfully made to overlap in $\mathbf{y}$ space. This is because each packet corresponds to a different particle velocity and to a different time of collision with the metal wall. Because the myriads of internal thermodynamic degrees of freedom are so chaotically complicated, it is very likely that as each of the n packets interacts with them, it will encounter different conditions, which will make the combined wave packet $f_n(\mathbf{y}, y_1, \cdots y_s)$ enter very different regions of $y_1, y_2\cdots y_s$ space. Thus, for all practical purposes, we can ignore the possibility that if two of the packets are made to cross in $\mathbf{y}$ space, the motion either of the atomic electron or of the outgoing particle will be affected.[18]

[17] See reference 2, Chapter 22, Sec. 11, for a treatment of a similar problem.

[18] It should be noted that exactly the same problem arises in the usual interpretation of the quantum theory for (reference 16), for whenever two packets overlap, then even in the usual interpretation, the system must be regarded as, in some sense, covering the states corresponding to both packets simultaneously. See reference 2, Chapter 6 and Chapter 16, Sec. 25. Once two packets have obtained classically describable separations, then, both in the usual interpretation and in our interpretation the probability that there will be significant interference between them is so overwhelmingly small that it may be compared to the probability that a tea kettle placed on a fire will happen to freeze instead of boil. Thus, we may for all practical purposes neglect the possibility of interference between packets corresponding to the different possible energy states in which the hydrogen atom may be left.

[19] See, for example, reference 2, Chapter 11, Sec. 17, and Chapter 12, Sec. 18.

8. PENETRATION OF A BARRIER

According to classical physics, a particle can never penetrate a potential barrier having a height greater than the particle kinetic energy. In the usual interpretation of the quantum theory, it is said to be able, with a small probability, to "leak" through the barrier. In our interpretation of the quantum theory, however, the potential provided by the Schroedinger ψ-field enables it to "ride" over the barrier, but only a few particles are likely to have trajectories that carry them all the way across without being turned around.

We shall merely sketch in general terms how the above results can be obtained. Since the motion of the particle is strongly affected by its ψ-field, we must first solve for this field with the aid of "Schroedinger's equation." Initially, we have a wave packet incident on the potential barrier; and because the probability density is equal to $|\psi(\mathbf{x})|^2$, the particle is certain to be somewhere within this wave packet. When the wave packet strikes the repulsive barrier, the ψ-field undergoes rapid changes which can be calculated[19] if desired, but whose precise form does not interest us here. At this time, the "quantum-mechanical" potential, $U=(-h^2/2m)\nabla^2R/R$, undergoes rapid and violent fluctuations, analogous to those described in Sec. 7 in connection with Eqs. (19), (20), and (25). The particle orbit then becomes very complicated and, because the potential is time dependent, very sensitive to the precise initial relationship between the particle position and the center of the wave packet. Ultimately, however, the incident wave packet disappears and is replaced by two packets, one of them a reflected packet and the other a transmitted packet having a much smaller intensity. Because the probability density is $|\psi|^2$, the particle must end up in one of these packets. The other packet can, as shown in Sec. 7, subsequently be ignored. Since the reflected packet is usually so much stronger than the transmitted packet, we conclude that during the time when the packet is inside the barrier, most of the particle orbits must be turned around, as a result of the violent fluctuations in the "quantum-mechanical" potential.

9. POSSIBLE MODIFICATIONS IN MATHEMATICAL FORMULATION LEADING TO EXPERIMENTAL PROOF THAT NEW INTERPRETATION IS NEEDED

We have already seen in a number of cases and in Paper II we shall prove in general, that as long as we

assume that ψ satisfies Schroedinger's equation, that $\mathbf{v}=\nabla S(\mathbf{x})/m$, and that we have a statistical ensemble with a probability density equal to $|\psi(\mathbf{x})|^2$, our interpretation of the quantum theory leads to physical results that are identical with those obtained from the usual interpretation. Evidence indicating the need for adopting our interpretation instead of the usual one could therefore come only from experiments, such as those involving phenomena associated with distances of the order of 10^{-13} cm or less, which are not now adequately understood in terms of the existing theory. In this paper we shall not, however, actually suggest any specific experimental methods of distinguishing between our interpretation and the usual one, but shall confine ourselves to demonstrating that such experiments are conceivable.

Now, there are an infinite number of ways of modifying the mathematical form of the theory that are consistent with our interpretation and not with the usual interpretation. We shall confine ourselves here, however, to suggesting two such modifications, which have already been indicated in Sec. 4, namely, to give up the assumption that $\mathbf{v}$ is necessarily equal to $\nabla S(\mathbf{x})/m$, and to give up the assumption that ψ must necessarily satisfy a homogeneous linear equation of the general type suggested by Schroedinger. As we shall see, giving up either of those first two assumptions will in general also require us to give up the assumption of a statistical ensemble of particles, with a probability density equal to $|\psi(\mathbf{x})|^2$.

We begin by noting that it is consistent with our interpretation to modify the equations of motion of a particle (8a) by adding any conceivable force term to the right-hand side. Let us, for the sake of illustration, consider a force that tends to make the difference, $\mathbf{p}-\nabla S(\mathbf{x})$, decay rapidly with time, with a mean decay time of the order of $\tau=10^{-13}/c$ seconds, where c is the velocity of light. To achieve this result, we write

$$m\frac{d^2\mathbf{x}}{dt^2}=-\nabla\left\{V(\mathbf{x})-\frac{\hbar^2}{2m}\frac{\nabla^2 R}{R}\right\}+\mathbf{f}(\mathbf{p}-\nabla S(\mathbf{x})), \quad (31)$$

where $\mathbf{f}(\mathbf{p}-\nabla S(\mathbf{x}))$ is assumed to be a function which vanishes when $\mathbf{p}=\nabla S(\mathbf{x})$ and more generally takes such a form that it implies a force tending to make $\mathbf{p}-\nabla S(\mathbf{x})$ decrease rapidly with the passage of time. It is clear, moreover, that f can be so chosen that it is large only in processes involving very short distances (where $\nabla S(\mathbf{x})$ should be large).

If the correct equations of motion resembled Eq. (31), then the usual interpretation would be applicable only over times much longer than τ, for only after such times have elapsed will the relation $\mathbf{p}=\nabla S(\mathbf{x})$ be a good approximation. Moreover, it is clear that such modifications of the theory cannot even be described in the usual interpretation, because they involve the precisely definable particle variables which are not postulated in the usual interpretation.

Let us now consider a modification that makes the equation governing ψ inhomogeneous. Such a modification is

$$i\hbar\partial\psi/\partial t=H\psi+\xi(\mathbf{p}-\nabla S(\mathbf{x}_i)). \quad (32)$$

Here, H is the usual Hamiltonian operator, $\mathbf{x}_i$ represents the actual location of the particle, and ξ is a function that vanishes when $\mathbf{p}=\nabla S(\mathbf{x}_i)$. Now, if the particle equations are chosen, as in Eq. (31), to make $\mathbf{p}-\nabla S(\mathbf{x}_i)$ decay rapidly with time, it follows that in atomic processes, the inhomogeneous term in Eq. (32) will become negligibly small, so that Schroedinger's equation is a good approximation. Nevertheless, in processes involving very short distances and very short times, the inhomogeneities would be important, and the ψ-field would, as in the case of the electromagnetic field, depend to some extent on the actual location of the particle.

It is clear that Eq. (32) is inconsistent with the usual interpretation of the theory. Moreover, we can contemplate further generalizations of Eq. (32), in the direction of introducing nonlinear terms that are large only for processes involving small distances. Since the usual interpretation is based on the hypothesis of linear superposition of "state vectors" in a Hilbert space, it follows that the usual interpretation could not be made consistent with such a nonlinear equation for a one-particle theory. In a many-particle theory, operators can be introduced, satisfying a nonlinear generalization of Schroedinger's equation; but these must ultimately operate on wave functions that satisfy a linear homogeneous Schroedinger equation.

Finally, we repeat a point already made in Sec. 4, namely, that if the theory is generalized in any of the ways indicated here, the probability density of particles will cease to equal $|\psi(\mathbf{x})|^2$. Thus, experiments would become conceivable that distinguish between $|\psi(\mathbf{x})|^2$ and this probability; and in this way we could obtain an experimental proof that the usual interpretation, which gives $|\psi(\mathbf{x})|^2$ *only* a probability interpretation, must be inadequate. Moreover, we shall show in Paper II that modifications like those suggested here would permit the particle position and momentum to be measured simultaneously, so that the uncertainty principle could be violated.

ACKNOWLEDGMENT

The author wishes to thank Dr. Einstein for several interesting and stimulating discussions.

A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II

DAVID BOHM*
Palmer Physical Laboratory, Princeton University
(Received July 5, 1951)

In this paper, we shall show how the theory of measurements is to be understood from the point of view of a physical interpretation of the quantum theory in terms of "hidden" variables, developed in a previous paper. We find that in principle, these "hidden" variables determine the precise results of each individual measurement process. In practice, however, in measurements that we now know how to carry out, the observing apparatus disturbs the observed system in an unpredictable and uncontrollable way, so that the uncertainty principle is obtained as a practical limitation on the possible precision of measurements. This limitation is not, however, inherent in the conceptual structure of our interpretation. We shall see, for example, that simultaneous measurements of position and momentum having unlimited precision would in principle be possible if, as suggested in the previous paper, the mathematical formulation of the quantum theory needs to be modified at very short distances in certain ways that are consistent with our interpretation but not with the usual interpretation.

We give a simple explanation of the origin of quantum-mechanical correlations of distant objects in the hypothetical experiment of Einstein, Podolsky, and Rosen, which was suggested by these authors as a criticism of the usual interpretation.

Finally, we show that von Neumann's proof that quantum theory is not consistent with hidden variables does not apply to our interpretation, because the hidden variables contemplated here depend both on the state of the measuring apparatus and the observed system and therefore go beyond certain of von Neumann's assumptions.

In two appendixes, we treat the problem of the electromagnetic field in our interpretation and answer certain additional objections which have arisen in the attempt to give a precise description for an individual system at the quantum level.

1. INTRODUCTION

IN a previous paper,[1] to which we shall hereafter refer as I, we have suggested an interpretation of the quantum theory in terms of "hidden" variables. We have shown that although this interpretation provides a conceptual framework that is broader than that of the usual interpretation, it permits of a set of three mutually consistent special assumptions, which lead to the same physical results as are obtained from the usual interpretation of the quantum theory. These three special assumptions are: (1) The ψ-field satisfies Schroedinger's equation. (2) If we write $\psi=R\exp(iS/\hbar)$, then the particle momentum is restricted to $\mathbf{p}=\nabla S(\mathbf{x})$. (3) We have a statistical ensemble of particle positions, with a probability density, $P=|\psi(\mathbf{x})|^2$. If the above three special assumptions are not made, then one obtains a more general theory that cannot be made consistent with the usual interpretation. It was suggested in Paper I that such generalizations may actually be needed for an understanding of phenomena associated with distances of the order of 10^{-13} cm or less, but may produce changes of negligible importance in the atomic domain.

In this paper, we shall apply the interpretation of the quantum theory suggested in Paper I to the development of a theory of measurements in order to show that as long as one makes the special assumptions indicated above, one is led to the same predictions for all measurements as are obtained from the usual interpretation. In our interpretation, however, the uncertainty principle is regarded, not as an inherent limitation on the precision with which we can correctly conceive of the simultaneous definition of momentum and position, but rather as a practical limitation on the precision with which these quantities can simultaneously be measured, arising from unpredictable and uncontrollable disturbances of the observed system by the measuring apparatus. If the theory needs to be generalized in the ways suggested in Paper I, Secs. 4 and 9, however, then these disturbances could in principle either be eliminated, or else be made subject to prediction and control, so that their effects could be corrected for. Our interpretation therefore demonstrates that measurements violating the uncertainty principle are at least conceivable.

2. QUANTUM THEORY OF MEASUREMENTS

We shall now show how the quantum theory of measurements is to be expressed in terms of our suggested interpretation of the quantum theory.[2]

In general, a measurement of any variable must always be carried out by means of an interaction of the system of interest with a suitable piece of measuring apparatus. The apparatus must be so constructed that any given state of the system of interest will lead to a certain range of states of the apparatus. Thus, the interaction introduces correlations between the state of the observed system and the state of the apparatus. The range of indefiniteness in this correlation may be called the uncertainty, or the error, in the measurement.

Let us now consider an observation designed to measure an arbitrary (hermitian) "observable" Q, associated with an electron. Let $\mathbf{x}$ represent the position of the electron, y that of the significant apparatus coordinate (or coordinates if there are more than one). Now, one can show[2] that it is enough to consider an impulsive measurement, i.e., a measurement utilizing a very strong interaction between apparatus and system

* Now at Universidade de São Paulo, Faculdade de Filosofia, Ciencias e Letras, São Paulo, Brasil.

[1] D. Bohm, Phys. Rev. **84**, 166 (1951).

[2] For a treatment of how the theory of measurements can be carried out with the usual interpretation, see D. Bohm, *Quantum Theory* (Prentice-Hall, Inc., New York, 1951), Chapter 22.

under observation, which lasts for so short a time that the changes of the apparatus and the system under observation that would have taken place in the absence of interaction can be neglected. Thus, at least while the interaction is taking place, we can neglect the parts of the Hamiltonian associated with the apparatus alone and with the observed system alone, and we need retain only the part of the Hamiltonian, H_I, representing the interaction. Moreover, if the Hamiltonian operator is chosen to be a function only of quantities that commute with Q, then the interaction process will produce no uncontrollable changes in the observable, Q, but only in observables that do not commute with Q. In order that the apparatus and the system under observation shall be coupled, however, it is necessary that H_I shall also depend on operators involving y.

For the sake of illustration of the principles involved, we shall consider the following interaction Hamiltonian:

$$H_I = -aQp_y, \tag{1}$$

where a is a suitable constant and p_y is the momentum conjugate to y.

Now, in our interpretation, the system is to be described by a four-dimensional but objectively real wave field that is a function of $\mathbf{x}$ and y and by a corresponding four-dimensional representative point, specified by the coordinates, $\mathbf{x}$, of the electron and the coordinate, y, of the apparatus. Since the motion of the representative point is in part determined by forces produced by the ψ-field acting on both electron and apparatus variables, our first step in solving this problem is to calculate the ψ-field. This is done by solving Schroedinger's equation, with the appropriate boundary conditions on ψ.

Now, during interaction, Schroedinger's equation is approximated by

$$i\hbar\partial\Psi/\partial t = -aQp_y\Psi = (ia/\hbar)Q\partial\Psi/\partial y. \tag{2}$$

It is now convenient to expand Ψ in terms of the complete set $\psi_q(\mathbf{x})$ of eigenfunctions of the operator, Q, where q denotes an eigenvalue of Q. For the sake of simplicity, we assume that the spectrum of Q is discrete, although the results are easily generalized to a continuous spectrum. Denoting the expansion coefficients by $f_q(y, t)$, we obtain

$$\Psi(\mathbf{x}, y, t) = \sum_q \psi_q(\mathbf{x}) f_q(y, t). \tag{3}$$

Noting that $Q\psi_q(\mathbf{x}) = q\psi_q(\mathbf{x})$, we readily verify that Eq. (2) can now be reduced to the following series of equations for $f_q(y, t)$:

$$i\hbar\partial f_q(y, t)/\partial t = (ia/\hbar)qf_q(y, t). \tag{4}$$

If the initial value of $f_q(y, t)$ was $f_q{}^0(y)$, we obtain as a solution

$$f_q(y, t) = f_q{}^0(y - aqt/\hbar^2), \tag{5}$$

and

$$\Psi(\mathbf{x}, y, t) = \sum_q \psi_q(\mathbf{x}) f_q{}^0(y - aqt/\hbar^2). \tag{6}$$

Now, initially the apparatus and the electron were independent. As shown in Paper I, Sec. 7, in our interpretation (as in the usual interpretation), independent systems must have wave fields $\Psi(\mathbf{x}, y, t)$ that are equal to a product of a function of $\mathbf{x}$ and a function of y. Initially, we therefore have

$$\Psi_0(\mathbf{x}, y) = \psi_0(\mathbf{x}) g_0(y) = g_0(y) \sum_q c_q \psi_q(\mathbf{x}), \tag{7}$$

where the c_q are the (unknown) expansion coefficients of $\psi_q(\mathbf{x})$, and $g_0(y)$ is the initial wave function of the apparatus coordinate, y. The function $g_0(y)$ will take the form of a packet. For the sake of convenience, we assume that this packet is centered at $y = 0$ and that its width is Δy. Normally, because the apparatus is classically describable, the definition of this packet is far less precise than that allowed by the limits of precision set by the uncertainty principle.

From Eqs. (7) and (3), we shall readily deduce that $f_q{}^0(y) = c_q g_0(y)$. When this value of $f_q{}^0(y)$ is inserted into Eq. (6), we obtain

$$\Psi(\mathbf{x}, y, t) = \sum_q c_q \psi_q(\mathbf{x}) g_0(y - aqt/\hbar^2). \tag{8}$$

Equation (8) indicates already that the interaction has introduced a correlation between q and the apparatus coordinate, y. In order to show what this correlation means in our interpretation of the quantum theory, we shall use some arguments that have been developed in more detail in Paper I, Sec. 7, in connection with a similar problem involving the interaction of two particles in a scattering process. First we note that while the electron and the apparatus are interacting, the wave function (8) becomes very complicated, so that if it is expressed as

$$\Psi(\mathbf{x}, y, t) = R(\mathbf{x}, y, t) \exp[iS(\mathbf{x}, y, t)/\hbar],$$

then R and S undergo rapid oscillations both as a function of position and of time. From this we deduce that the "quantum-mechanical" potential,

$$U = (-\hbar^2/2mR)(\nabla_x{}^2 R + \partial^2 R/\partial y^2),$$

undergoes violent fluctuations, especially where R is small, and that the particle momenta, $\mathbf{p} = \nabla_x S(\mathbf{x}, y, t)$ and $p_y = \partial S(\mathbf{x}, y, t)/\partial y$, also undergo corresponding violent and extremely complicated fluctuations. Eventually, however, if the interaction continues long enough, the behavior of the system will become simpler because the packets $g_0(y - aqt/\hbar^2)$, corresponding to different values of q, will cease to overlap in y space. To prove this, we note that the center of the qth packet in y space is at

$$y = aqt/\hbar^2; \quad \text{or} \quad q = \hbar^2 y/at. \tag{9}$$

If we denote the separation of adjacent values of q by δq, we then obtain for the separation of the centers of adjacent packets in y space

$$\delta y = at\delta q/\hbar^2. \tag{10}$$

It is clear that if the product of the strength of interaction a, and the duration of interaction, t, is large enough, then δy can be made much larger than the

width Δy of the packet. Then packets corresponding to different values of q will cease to overlap in y space and will, in fact, obtain separations large enough to be classically describable.

Because the probability density is equal to $|\Psi|^2$, we deduce that the apparatus variable, y, must finally enter one of the packets and remain with that packet thereafter (since it does not enter the intermediate space between packets in which the probability density is practically zero). Now, the packet entered by the apparatus variable y determines the actual result of the measurement, which the observer will obtain when he looks at the apparatus. The other packets can (as shown in Paper I, Sec. 7) be ignored, because they affect neither the quantum-mechanical potential acting on the particle coordinates $\mathbf{x}$ and y, nor the particle momenta, $\mathbf{p}_x=\nabla_x S$ and $p_y=\partial S/\partial y$. Moreover, the wave function can also be renormalized without affecting any of the above quantities. Thus, for all practical purposes, we can replace the complete wave function, Eq. (8), by a new renormalized wave function

$$\Psi(\mathbf{x}, y)=\psi_q(\mathbf{x})g_0(y-aqt/\hbar^2), \tag{11}$$

where q now corresponds to the packet actually containing the apparatus variable, y. From this wave function, we can deduce, as shown in Paper I, Sec. 7, that the apparatus and the electron will subsequently behave independently. Moreover, by observing the approximate value of the apparatus coordinate within an error $\Delta y \ll \delta y$, we can deduce with the aid of Eq. (9) that since the electron wave function can for all practical purposes be regarded as $\psi_q(\mathbf{x})$, the observable, Q, must have the definite value, q. However, if the product, "$at\delta q/\hbar^2$," appearing in Eqs. (8), (9), (10), and (11), had been less than Δy, then no clear measurement of Q would have been possible, because packets corresponding to different q would have overlapped, and the measurement would not have had the requisite accuracy.[3]

Finally, we note that even if the apparatus packets are subsequently caused to overlap, none of those conclusions will be altered. For the apparatus variable y will inevitably be coupled to a whole host of internal thermodynamic degrees of freedom, $y_1, y_2, \cdots y_n$, as a result of effects such as friction and brownian motion. As shown in Paper I, Sec. 7, interference between packets corresponding to different values of q would be possible only if the packets overlapped in the space of $y_1, y_2, \cdots y_n$, as well as in y space. Such an overlap, however, is so improbable that for all practical purposes, we can ignore the possibility that it will ever occur.

3. THE ROLE OF PROBABILITY IN MEASUREMENTS—THE UNCERTAINTY PRINCIPLE

In principle, the final result of a measurement is determined by the initial form of the wave function of the combined system, $\Psi^0(\mathbf{x}, y)$, and by the initial position of the electron particle, $\mathbf{x}_0$, and the apparatus variable, y_0. In practice, however, as we have seen, the orbit fluctuates violently while interaction takes place, and is very sensitive to the precise initial values of $\mathbf{x}$ and y, which we can neither predict nor control. All that we can predict in practice is that in an ensemble of similar experiments performed under equivalent initial conditions, the probability density is $|\Psi(\mathbf{x}, y)|^2$. From this information, however, we are able to calculate only the probability that in an individual experiment, the result of a measurement of Q will be a specific number q. To obtain the probability of a given value of q, we need only integrate the above probability density over all $\mathbf{x}$ and over all values of y in the neighborhood of the qth packet. Because the packets do not overlap, the Ψ-field in this region is equal to $c_q\psi_q(\mathbf{x})g_0(y-aqt/\hbar^2)$ [see Eq. (8)]. Since, by definition, $\psi_q(\mathbf{x})$ and $g_0(y)$ are normalized, the total probability that a particle is in the qth packet is

$$P_q=|c_q|^2. \tag{12}$$

The above is, however, just what is obtained from the usual interpretation. We conclude then that our interpretation is capable of leading in all possible experiments to identical predictions with those obtained from the usual interpretation (provided, or course, that we make the special assumptions indicated in the introduction).

Let us now see what a measurement of the observable, Q, implies with regard to the state of the electron particle and its Ψ-field. First, we note that the process of interaction with an apparatus designed to measure the observable, Q, effectively transforms the electron ψ-field from whatever it was before the measurement took place into an eigenfunction $\psi_q(\mathrm{x})$ of the operator Q. The precise value of q that comes out of this process is as we have seen, not, in general, completely predictable or controllable. If, however, the same measurement is repeated after the ψ-field has been transformed into $\psi_q(\mathbf{x})$, we can then predict that (as in the usual interpretation), the same value of q, and therefore the same wave function, $\psi_q(\mathbf{x})$, will be obtained again. If, however, we measure an observable "P" that does not commute with Q, then the results of this measurement are not, in practice, predictable or controllable. For as shown in Eq. (8), the Ψ-field after interaction with the measuring apparatus is now transformed into

$$\Psi(\mathbf{x}, z, t)=\textstyle\sum_p a_{p,q}\phi_p(\mathbf{x})g_0(z-apt/\hbar^2), \tag{13}$$

where $\phi_p(\mathbf{x})$ is an eigenfunction of the operator, P, belonging to an eigenvalue, p, and where $a_{p,q}$ is an expansion coefficient defined by

$$\psi_q(\mathbf{x})=\textstyle\sum_p a_{p,q}\phi_p(\mathbf{x}). \tag{14}$$

Since the packets corresponding to different p ultimately become completely separate in z space, we deduce, as in the case of the measurement of Q, that

[3] A similar requirement is obtained in the usual interpretation. See reference 2, Chapter 22, Sec. 8.

for all practical purposes, this wave function may be replaced by

$$\Psi = a_{pq}\phi_p(\mathbf{x})g_0(z - apt/\hbar^2),$$

where p now represents the packet actually entered by the apparatus coordinate, y. As in the case of measurement of Q, we readily show that the precise value of p that comes out of this experiment cannot be predicted or controlled and that the probability of a given value of p is equal to $|a_{pq}|^2$. This is, however, just what is obtained in the usual interpretation of this process.

It is clear that if two "observables," P and Q, do not commute, one cannot carry out a measurement of both simultaneously on the same system. The reason is that each measurement disturbs the system in a way that is incompatible with carrying out the process necessary for the measurement of the other. Thus, a measurement of P requires that wave field, ψ, shall become an eigenfunction of P, while a measurement of Q requires that it shall become an eigenfunction of Q. If P and Q do not commute, then by definition, no ψ-function can be simultaneously an eigenfunction of both. In this way, we understand in our interpretation why measurements, of complementary quantities, must (as in the usual interpretation) necessarily be limited in their precision by the uncertainty principle.

4. PARTICLE POSITIONS AND MOMENTA AS "HIDDEN VARIABLES"

We have seen that in measurements that can now be carried out, we cannot make precise inferences about the particle position, but can say only that the particle must be somewhere in the region in which $|\psi|$ is appreciable. Similarly, the momentum of a particle that happens to be at the point, $\mathbf{x}$, is given by $\mathbf{p} = \nabla S(\mathbf{x})$, so that since $\mathbf{x}$ is not known, the precise value of $\mathbf{p}$ is also not, in general, inferrable. Hence, as long as we are restricted to making observations of this kind, the precise values of the particle position and momentum must, in general, be regarded as "hidden," since we cannot at present measure them. They are, however, connected with real and already observable properties of matter because (along with the ψ-field) they determine in principle the actual result of each individual measurement. By way of contrast, we recall here that in the usual interpretation of the theory, it is stated that although each measurement admittedly leads to a definite number, nothing determines the actual value of this number. The result of each measurement is assumed to arise somehow in an inherently indescribable way that is not subject to a detailed analysis. Only the statistical results are said to be predictable. In our interpretation, however, we assert that the at present "hidden" precisely definable particle positions and momenta determine the results of each individual measurement process, but in a way whose precise details are so complicated and uncontrollable, and so little known, that one must for all practical purposes restrict oneself to a statistical description of the connection between the values of these variables and the directly observable results of measurements. Thus, we are unable at present to obtain direct experimental evidence for the existence of precisely definable particle positions and momenta.

5. "OBSERVABLES" OF USUAL INTERPRETATION ARE NOT A COMPLETE DESCRIPTION OF SYSTEM IN OUR INTERPRETATION

We have seen in Sec. 3 that in the measurement of an "observable," Q, we cannot obtain enough information to provide a complete specification of the state of an electron, because we cannot infer the precisely defined values of the particle momentum and position, which are, for example, needed if we wish to make precise predictions about the future behavior of the electron. Moreover, the process of measuring an observable does not provide any unambiguous information about the state that existed before the measurement took place; for in such a measurement, the ψ-field is transformed into an in practice unpredictable and uncontrollable eigenfunction, $\psi_q(\mathbf{x})$, of the measured "observable" Q. This means that the measurement of an "observable" is not really a measurement of any physical property belonging to the observed system alone. Instead, the value of an "observable" measures only an incompletely predictable and controllable potentiality belonging just as much to the measuring apparatus as to the observed system itself.[4] At best, such a measurement provides unambiguous information only at a classical level of accuracy, where the disturbance of the ψ-field by the measuring apparatus can be neglected. The usual "observables" are therefore not what we ought to try to measure at a quantum level of accuracy. In Sec. 6, we shall see that it is conceivable that we may be able to carry out new kinds of measurements, providing information not about "observables" having a very ambiguous significance, but rather about physically significant properties of a system, such as the actual values of the particle position and momentum.

As an example of the rather indirect and ambiguous significance of the "observable," we may consider the problem of measuring the momentum of an electron. Now, in the usual interpretation, it is stated that one can always measure the momentum "observable" without changing the value of the momentum. The result is said, for example, to be obtainable with the aid of an impulsive interaction involving only operators which commute with the momentum operator, $\mathbf{p}_x$. To represent such a measurement, we could choose $H_I = -ap_xp_y$ in Eq. (1). In our interpretation, however, we cannot in general conclude that such an interaction will enable us to measure the actual particle momentum without changing its value. In fact, in our interpreta-

[4] Even in the usual interpretation, an observation must be regarded as yielding a measure of such a potentiality. See reference 2, Chapter 6, Sec. 9.

tion, a measurement of particle momentum that does not change the value of this momentum is possible only if the ψ-field initially takes the special form, $\exp(i\mathbf{p}\cdot\mathbf{x}/\hbar)$. If, however, ψ initially takes its most general possible form,

$$\psi=\sum_p a_1\mathbf{p}\exp(i\mathbf{p}\cdot\mathbf{x}/\hbar), \tag{15}$$

then as we have seen in Secs. 2 and 3, the process of measuring the "observable" p_x will effectively transform the ψ-field of the electron into

$$\exp(ip_x/\hbar) \tag{16}$$

with a probability $|a_p|^2$ that a given value of p_x will be obtained. When the ψ-field is altered in this way, large quantities of momentum can be transferred to the particle by the changing ψ-field, even though the interaction Hamiltonian, H_I, commutes with the momentum operator, $\mathbf{p}$.

As an example, we may consider a stationary state of an atom, of zero angular momentum. As shown in Paper I, Sec. 5, the ψ-field for such a state is real, so that we obtain

$$\mathbf{p}=\nabla S=0.$$

Thus, the particle is at rest. Nevertheless, we see from Eqs. (14) and (15) that if the momentum "observable" is measured, a large value of this "observable" may be obtained if the ψ-field happens to have a large fourier coefficient, a_p, for a high value of $\mathbf{p}$. The reason is that in the process of interaction with the measuring apparatus, the ψ-field is altered in such a way that it can give the electron particle a correspondingly large momentum, thus transforming some of the potential energy of interaction of the particle with its ψ-field into kinetic energy.

A more striking illustration of the points discussed above is afforded by the problem of a "free" particle contained between two impenetrable and perfectly reflecting walls, separated by a distance L. For this case, the spatial part of the ψ-field is

$$\psi=\sin(2\pi nx/L),$$

where n is an integer and the energy of the electron is

$$E=(1/2m)(nh/L)^2.$$

Because the ψ-field is real, we deduce that the particle is at rest.

Now, at first sight, it may seem puzzling that a particle having a high energy should be at rest in the empty space between two walls. Let us recall, however, that the space is not really empty, but contains an objectively real ψ-field that can act on the particle. Such an action is analogous to (but of course not identical with) the action of an electromagnetic field, which could create non-uniform motion of the particle in this apparently "empty" enclosure. We observe that in our problem, the ψ-field is able to bring the particle to rest and to transform the entire kinetic energy into potential energy of interaction with the ψ-field. To prove this, we evaluate the "quantum-mechanical potential" for this ψ-field

$$U=\frac{-\hbar^2\nabla^2R}{2m}\frac{}{R}=\frac{-\hbar^2\nabla^2\psi}{2m}\frac{}{\psi}=\frac{1}{2m}\left(\frac{nh}{L}\right)^2$$

and note that it is precisely equal to the total energy, E.

Now, as we have seen, any measurement of the momentum "observable" must change the ψ-field in such a way that in general some (and in our case, all) of this potential energy is transformed into kinetic energy. We may use as an illustration of this general result a very simple specific method of measuring the momentum "observable," namely, to remove the confining walls suddenly and then to measure the distance moved by the particle after a fairly long time. We can compute the momentum by dividing this distance by the time of transit. If (as in the usual interpretation of the quantum theory) we assume that the electron is "free," then we conclude that the process of removing the walls should not appreciably change the momentum if we do it fast enough, for the probability that the particle is near a wall when this happens can then in principle be made arbitrarily small. In our interpretation, however, the removal of the walls alters the particle momentum indirectly, because of its effect on the ψ-field, which acts on the particle. Thus, after the walls are removed, two wave packets moving in opposite directions begin to form, and ultimately they become completely separate in space. Because the probability density is $|\psi|^2$, we deduce that the particle must end up in one packet or the other. Moreover, the reader will readily convince himself that the particle momentum will be very close to $\pm nh/L$, the sign depending on which packet the particle actually enters. As in Sec. (2), the packet not containing the particle can subsequently be ignored. In principle, the final particle momentum is determined by the initial form of the ψ-field and by the initial particle position. Since we do not in practice know the latter, we can at best predict a probability of $\frac{1}{2}$ that the particle ends up in either packet. We conclude then that this measurement of the momentum "observable" leads to the same result as is predicted in the usual interpretation. However, the actual particle momentum existing before the measurement took place is quite different from the numerical value obtained for the momentum "observable," which, in the usual interpretation, is called the "momentum."

6. ON THE POSSIBILITY OF MEASUREMENTS OF UNLIMITED PRECISION

We have seen that the so-called "observables" do not measure any very readily interpretable properties of a system. For example, the momentum "observable" has in general no simple relation to the actual particle momentum. It may therefore be fruitful to consider

how we might try to measure properties which, according to our interpretation, are (along with the ψ-field) the physically significant properties of an electron, namely, the actual particle position and momentum. In connection with this problem, we shall show that if, as suggested in Paper I, Secs. 4 and 9, we give up the three mutually consistent special assumptions leading to the same results as those of the usual interpretation of the quantum theory, then in our interpretation, the particle position and momentum can in principle be measured simultaneously with unlimited precision.

Now, for our purposes, it will be adequate to show that precise predictions of the future behavior of a system are in principle possible. In our interpretation, a sufficient condition for precise predictions is as we have seen that we shall be able to prepare a system in a state in which the ψ-field and the initial particle position and momentum are precisely known. We have shown that it is possible, by measuring the "observable," Q, with the aid of methods that are now available, to prepare a state in which the ψ-field is effectively transformed into a known form, $\psi_q(\mathbf{x})$; but we cannot in general predict or control the precise position and momentum of the particle. If we could now measure the position and momentum of the particle without altering the ψ-field, then precise predictions would be possible. However, the results of Secs. 2, 3, and 4 prove that as long as the three special assumptions indicated above are valid, we cannot measure the particle position more accurately without effectively transforming the ψ-function into an incompletely predictable and controllable packet that is much more localized than $\psi_q(\mathbf{x})$. Thus, efforts to obtain more precise definition of the state of the system will be defeated. But it is clear that the difficulty originates in the circumstance that the potential energy of interaction between electron and apparatus, $V(\mathbf{x}, y)$, plays two roles. For it not only introduces a direct interaction between the two particles, proportional in strength to $V(\mathbf{x}, y)$ itself, but it introduces an indirect interaction between these particles, because this potential also appears in the equation governing the ψ-field. This indirect interaction may involve rapid and violent fluctuations, even when $V(\mathbf{x}, y)$ is small. Thus, we are led to lose control of the effects of this interaction, because no matter how small $V(\mathbf{x}, y)$ is, very large and chaotically complicated disturbances in the particle motion may occur.

If, however, we give up the three special assumptions mentioned previously, then it is not inherent in our conceptual structure that every interaction between particles must inevitably also produce large and uncontrollable changes in the ψ-field. Thus, in Paper I, Eq. (31), we give an example in which we postulate a force acting on a particle that is not necessarily accompanied by a corresponding change in the ψ-field. Equation (31), Paper I, is concerned only with a one-particle system, but similar assumptions can be made for systems of two or more particles. In the absence of any specific theory, our interpretation permits an infinite number of kinds of such modifications, which can be chosen to be important at small distances but negligible in the atomic domain. For the sake of illustration, suppose that it should turn out that in certain processes connected with very small distances, the force acting on the apparatus variable is

$$F_y = ax,$$

where a is a constant. Now if "a" is made large enough so that the interaction is impulsive, we can neglect all changes in y that are brought about by the forces that would have been present in the absence of this interaction. Moreover, for the sake of illustration of the principles involved, we are permitted to make the assumption, consistent with our interpretation, that the force on the electron is zero. The equation of motion of y is then

$$\ddot{y} = ax/m.$$

The solution is

$$y - y_0 = (axt^2/2m) + \dot{y}_0 t,$$

where $\dot{y}_0$ is the initial velocity of the apparatus variable and y_0 its initial position. Now, if the product, at^2, is large enough, then $y - y_0$ can be made much larger than the uncertainty in y arising from the uncertainty of y_0, and the uncertainty of $\dot{y}_0$. Thus, $y - y_0$ will be determined primarily by the particle position, x. In this way, it is conceivable that we could obtain a measurement of x that does not significantly change $x, \dot{x}$, or the ψ-function. The particle momentum can then be obtained from the relation, $p = \nabla S(\mathbf{x})$, where $S/\hbar$ is the phase of the ψ-function. Thus, precise predictions would in principle be possible.

7. THE ORIGIN OF THE STATISTICAL ENSEMBLE IN THE QUANTUM THEORY

We shall now see that even if, because of a failure of the three special assumptions mentioned in Secs. 1 and 6, we are able to determine the particle positions and momenta precisely, we shall nevertheless ultimately obtain a statistical ensemble again at the atomic level, with a probability density equal to $|\psi|^2$. The need for such an ensemble arises from the chaotically complicated character of the coupling between the electron and classical systems, such as volumes of gas, walls of containers, pieces of measuring apparatus, etc., with which this particle must inevitably in practice interact. For as we have seen in Sec. 2, and in Paper I, Sec. 7, during the course of such an interaction, the "quantum-mechanical" potential undergoes violent and rapid fluctuations, which tend to make the particle orbit wander over the whole region in which the ψ-field is appreciable. Moreover, these fluctuations are further complicated by the effects of molecular chaos in the very large number of internal thermodynamic degrees of freedom of these classically describable systems, which are inevitably excited in any interaction process.

Thus, even if the initial particle variables were well defined, we should soon in practice lose all possibility of following the particle motion and would be forced to have recourse to some kind of statistical theory. The only question that remains is to show why the probability density that ultimately comes about should be equal to $|\psi|^2$ and not to some other quantity.

To answer this question, we first note that a statistical ensemble with a probability density $|\psi(\mathbf{x})|^2$ has the property that under the action of forces which prevail at the atomic level, where our three special assumptions are satisfied, it will be preserved by the equations of motion of the particles, once it comes into existence. There remains only the problem of showing that an arbitrary deviation from this ensemble tends, under the action of the chaotically complicated forces described in the previous paragraph, to decay into an ensemble with a probability density of $|\psi(\mathbf{x})|^2$. This problem is very similar to that of proving Boltzmann's H theorem, which shows in connection with a different but analogous problem that an arbitrary ensemble tends as a result of molecular chaos to decay into an equilibrium Gibbs ensemble. We shall not carry out a detailed proof here, but we merely suggest that it seems plausible that one could along similar lines prove that in our problem, an arbitrary ensemble tends to decay into an ensemble with a density of $|\psi(\mathbf{x})|^2$. These arguments indicate that in our interpretation, quantum fluctuations and classical fluctuations (such as the Brownian motion) have basically the same origin; *viz.*, the chaotically complicated character of motion at the microscopic level.

8. THE HYPOTHETICAL EXPERIMENT OF EINSTEIN, PODOLSKY, AND ROSEN

The hypothetical experiment of Einstein, Podolsky, and Rosen[8] is based on the fact that if we have two particles, the sum of their momenta, $p=p_1+p_2$, commutes with the difference of their positions, $\xi=x_1-x_2$. We can therefore define a wave function in which p is zero, while ξ has a given value, a. Such a wave function is

$$\psi=\delta(x_1-x_2-a). \tag{17}$$

In the usual interpretation of the quantum theory, p_1-p_2 and x_1+x_2 are completely undetermined in a system having the above wave function.

The whole experiment centers on the fact that an observer has a choice of measuring either the momentum or the position of any one of the two particles. Whichever of these quantities he measures, he will be able to infer a definite value of the corresponding variable in the other particle, because of the fact that the above wave function implies correlations between variables belonging to each particle. Thus, if he obtains a position x_1 for the first particle, he can infer a position of $x_2=a-x_1$ for the second particle; but he loses all possibility of making any inferences about the momenta of either particle. On the other hand, if he measures the momentum of the first particle and obtains a value of p_1, he can infer a value of $p_2=-p_1$ for the momentum of the second particle; but he loses all possibility of making any inferences about the position of either particle. Now, Einstein, Podolsky, and Rosen believe that this result is itself probably correct, but they do not believe that quantum theory as usually interpreted can give a complete description of how these correlations are propagated. Thus, if these were classical particles, we could easily understand the propagation of correlations because each particle would then simply move with a velocity opposite to that of the other. But in the usual interpretation of quantum theory, there is no similar conceptual model showing in detail how the second particle, which is not in any way supposed to interact with the first particle, is nevertheless able to obtain either an uncontrollable disturbance of its position or an uncontrollable disturbance of its momentum depending on what kind of measurement the observer decided to carry out on the first particle. Bohr's point of view is, however, that no such model should be sought and that we should merely accept the fact that these correlations somehow manage to appear. We must note, of course, that the quantum-mechanical description of these processes will always be consistent, even though it gives us no precisely definable means of describing and analyzing the relationships between the classically describable phenomena appearing in various pieces of measuring apparatus.

In our suggested new interpretation of the quantum theory, however, we can describe this experiment in terms of a single precisely definable conceptual model, for we now describe the system in terms of a combination of a six-dimensional wave field and a precisely definable trajectory in a six-dimensional space (see Paper I, Sec. 6). If the wave function is initially equal to Eq. (17), then since the phase vanishes, the particles are both at rest. Their possible positions are, however, described by an ensemble, in which $x_1-x_2=a$. Now, if we measure the position of the first particle, we introduce uncontrollable fluctuations in the wave function for the entire system, which, through the "quantum-mechanical" forces, bring about corresponding uncontrollable fluctuations in the momentum of each particle. Similarly, if we measure the momentum of the first particle, uncontrollable fluctuations in the wave function for the system bring about, through the "quantum-mechanical" forces, corresponding uncontrollable changes in the position of each particle. Thus, the "quantum-mechanical" forces may be said to transmit uncontrollable disturbances instantaneously from one particle to another through the medium of the ψ-field.

What does this transmission of forces at an infinite rate mean? In nonrelativistic theory, it certainly causes

[8] Einstein, Podolsky, and Rosen, Phys. Rev. **47**, 777 (1933).

no difficulties. In a relativistic theory, however, the problem is more complicated. We first note that as long as the three special assumptions mentioned in Sec. 2 are valid, our interpretation can give rise to no inconsistencies with relativity, because it leads to precisely the same predictions for all physical processes as are obtained from the usual interpretation (which is known to be consistent with relativity). The reason why no contradictions with relativity arise in our interpretation despite the instantaneous transmission of momentum between particles is that no signal can be carried in this way. For such a transmission of momentum could constitute a signal only if there were some practical means of determining precisely what the second particle would have done if the first particle had not been observed; and as we have seen, this information cannot be obtained as long as the present form of the quantum theory is valid. To obtain such information, we require conditions (such as might perhaps exist in connection with distances of the order of 10^{-13} cm) under which the usual form of the quantum theory breaks down (see Sec. 6), so that the positions and momenta of the particles can be determined simultaneously and precisely. If such conditions should exist, then there are two ways in which contradictions might be avoided. First, the more general physical laws appropriate to the new domains may be such that they do not permit the transmission of controllable aspects of interparticle forces faster than light. In this way, Lorentz covariance could be preserved. Secondly, it is possible that the application of the usual criteria of Lorentz covariance may not be appropriate when the usual interpretation of quantum theory breaks down. Even in connection with gravitational theory, general relativity indicates that the limitation of speeds to the velocity of light does not necessarily hold universally. If we adopt the spirit of general relativity, which is to seek to make the properties of space dependent on the properties of the matter that moves in this space, then it is quite conceivable that the metric, and therefore the limiting velocity, may depend on the ψ-field as well as on the gravitational tensor $g^{\mu,\nu}$. In the classical limit, the dependence on the ψ-field could be neglected, and we would get the usual form of covariance. In any case, it can hardly be said that we have a solid experimental basis for requiring the same form of covariance at very short distances that we require at ordinary distances.

To sum up, we may assert that wherever the present form of the quantum theory is correct, our interpretation cannot lead to inconsistencies with relativity. In the domains where the present theory breaks down, there are several possible ways in which our interpretation could continue to treat the problem of covariance consistently. The attempt to maintain a consistent treatment of covariance in this problem might perhaps serve as an important heuristic principle in the search for new physical laws.

9. ON VON NEUMANN'S DEMONSTRATION THAT QUANTUM THEORY IS INCONSISTENT WITH HIDDEN VARIABLES

Von Neumann[6] has studied the following question: "If the present mathematical formulation of the quantum theory and its usual probability intepretation are assumed to lead to absolutely correct results for every experiment that can ever be done, can quantum-mechanical probabilities be understood in terms of any conceivable distribution over hidden parameters?" Von Neumann answers this question in the negative. His conclusions are subject, however, to the criticism that in his proof he has implicitly restricted himself to an excessively narrow class of hidden parameters and in this way has excluded from consideration precisely those types of hidden parameters which have been proposed in this paper.

To demonstrate the above statements, we summarize Von Neumann's proof briefly. This proof (which begins on p. 167 of his book), shows that the usual quantum-mechanical rules of calculating probabilities imply that there can be no "dispersionless states," i.e., states in which the values of all possible observables are simultaneously determined by physical parameters associated with the observed system. For example, if we consider two noncommuting observables, p and q, then Von Neumann shows that it would be inconsistent with the usual rules of calculating quantum-mechanical probabilities to assume that there were in the observed system a set of hidden parameters which simultaneously determined the results of measurements of position and momentum "observables." With this conclusion, we are in agreement. However, in our suggested new interpretation of the theory, the so-called "observables" are, as we have seen in Sec. 5, not properties belonging to the observed system alone, but instead potentialities whose precise development depends just as much on the observing apparatus as on the observed system. In fact, when we measure the momentum "observable," the final result is determined by hidden parameters in the momentum-measuring device as well as by hidden parameters in the observed electron. Similarly, when we measure the position "observable," the final result is determined in part by hidden parameters in the position-measuring device. Thus, the statistical distribution of "hidden" parameters to be used in calculating averages in a momentum measurement is different from the distribution to be used in calculating averages in a position measurement. Von Neumann's proof (see p. 171 in his book) that no single distribution of hidden parameters could be consistent with the results of the quantum theory is therefore irrelevant here, since in our interpretation of measurements of the type that can now be carried out, the distribution of hidden parameters varies in accordance with the different mutually exclusive experimental

[6] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Verlag. Julius Springer, Berlin, 1932).

arrangements of matter that must be used in making different kinds of measurements. In this point, we are in agreement with Bohr, who repeatedly stresses the fundamental role of the measuring apparatus as an inseparable part of the observed system. We differ from Bohr, however, in that we have proposed a method by which the role of the apparatus can be analyzed and described in principle in a precise way, whereas Bohr asserts that a precise conception of the details of the measurement process is as a matter of principle unattainable.

Finally, we wish to stress that the conclusions drawn thus far refer only to the measurement of the so-called "observables" carried out by the methods that are now available. If the quantum theory needs to be modified at small distances, then, as suggested in Sec. 6, precise measurements can in principle be made of the actual position and momentum of a particle. Here, it should be noted that Von Neumann's theorem is likewise irrelevant, this time because we are going beyond the assumption of the unlimited validity of the present general form of quantum theory, which plays an integral part in his proof.

10. SUMMARY AND CONCLUSIONS

The usual interpretation of the quantum theory implies that we must renounce the possibility of describing an individual system in terms of a single precisely defined conceptual model. We have, however, proposed an alternative interpretation which does not imply such a renunciation, but which instead leads us to regard a quantum-mechanical system as a synthesis of a precisely definable particle and a precisely definable ψ-field which exerts a force on this particle. An experimental choice between these two interpretations cannot be made in a domain in which the present mathematical formulation of the quantum theory is a good approximation; but such a choice is conceivable in domains, such as those associated with dimensions of the order of 10^{-13} cm, where the extrapolation of the present theory seems to break down and where our suggested new interpretation can lead to completely different kinds of predictions.

At present, our suggested new interpretation provides a consistent alternative to the usual assumption that no objective and precisely definable description of reality is possible at the quantum level of accuracy. For, in our description, the problem of objective reality at the quantum level is at least in principle not fundamentally different from that at the classical level, although new problems of measurement of the properties of an individual system appear, which can be solved only with the aid of an improvement in the theory, such as the possible modifications in the nuclear domain suggested in Sec. 6. In this connection, we wish to point out that what we can measure depends not only on the type of apparatus that is available, but also on the existing theory, which determines the kind of inference that can be used to connect the directly observable state of the apparatus with the state of the system of interest. In other words, our epistemology is determined to a large extent by the existing theory. It is therefore not wise to specify the possible forms of future theories in terms of purely epistomological limitations deduced from existing theories.

The development of the usual interpretation of the quantum theory seems to have been guided to a considerable extent by the principle of not postulating the possible existence of entities which cannot now be observed. This principle, which stems from a general philosophical point of view known during the nineteenth century as "positivism" or "empiricism" represents an extraphysical limitation on the possible kinds of theories that we shall choose to take into consideration.[7] The word "extraphysical" is used here advisedly, since we can in no way deduce, either from the experimental data of physics, or from its mathematical formulation, that it will necessarily remain forever impossible for us to observe entities whose existence cannot now be observed. Now, there is no reason why an extraphysical general principle is necessarily to be avoided, since such principles could conceivably serve as useful working hypotheses. The particular extraphysical principle described above cannot, however, be said to be a good working hypothesis. For the history of scientific research is full of examples in which it was very fruitful indeed to assume that certain objects or elements might be real, long before any procedures were known which would permit them to be observed directly. The atomic theory is just such an example. For the possibility of the actual existence of individual atoms was first postulated in order to explain various macrophysical results which could, however, also be understood directly in terms of macrophysical concepts without the need for assuming the existence of atoms. Certain nineteenth-century positivists (notably Mach) therefore insisted on purely philosophical grounds that it was incorrect to suppose that individual atoms actually existed, because they had never been observed as such. The atomic theory, they thought, should be regarded only as an interesting way of calculating various observable large-scale properties of matter. Nevertheless, evidence for the existence of individual atoms was ultimately discovered by people who took the atomic hypothesis seriously enough to suppose that individual atoms might actually exist, even though no one had yet observed them. We may have here, perhaps, a close analogy to the usual interpretation of the quantum theory, which avoids considering the possibility that the wave function of an individual system may represent objective reality, because we cannot observe it with the aid of existing experiments and theories.

[7] A leading nineteenth-century exponent of the positivist point of view was Mach. Modern positivists appear to have retreated from this extreme position, but its reflection still remains in the philosophical point of view implicitly adopted by a large number of modern theoretical physicists.

Finally, as an alternative to the positivist hypothesis of assigning reality only to that which we can now observe, we wish to prevent here another hypothesis, which we believe corresponds more closely to conclusions that can be drawn from general experience in actual scientific research. This hypothesis is based on the simple assumption that the world as a whole is objectively real and that, as far as we now know, it can correctly be regarded as having a precisely describable and analyzable structure of unlimited complexity. The pattern of this structure seems to be reflected completely but indirectly at every level, so that from experiments done at the level of size of human beings, it is very probably possible ultimately to draw inferences concerning the properties of the whole structure at all levels. We should never expect to obtain a complete theory of this structure, because there are almost certainly more elements in existence than we possibly can be aware of at any particular stage of scientific development. Any specified element, however, can in principle ultimately be discovered, but never all of them. Of course, we must avoid postulating a new element for each new phenomenon. But an equally serious mistake is to admit into the theory only those elements which can now be observed. For the purpose of a theory is not only to correlate the results of observations that we already know how to make, but also to suggest the need for new kinds of observations and to predict their results. In fact, the better a theory is able to suggest the need for new kinds of observations and to predict their results correctly, the more confidence we have that this theory is likely to be good representation of the actual properties of matter and not simply an empirical system especially chosen in such a way as to correlate a group of already known facts.

APPENDIX A. PHOTOELECTRIC AND COMPTON EFFECTS

In this appendix, we shall show how the electromagnetic field is to be described in our new interpretation, with the purpose of making possible a treatment of the photoelectric and Compton effects. For our purposes, it is adequate to restrict ourselves to a gauge in which $\mathrm{div}\mathbf{A}=0$, and to consider only the transverse part of the electromagnetic field, for in this gauge, the longitudinal part of the field can be expressed through Poisson's equation entirely in terms of the charge density. The Fourier analysis of the vector potential is then

$$\mathbf{A}(\mathbf{x})=(4\pi/V)^{\frac{1}{2}}\sum_{\mathbf{k},\mu}\boldsymbol{\epsilon}_{\mathbf{k},\mu}q_{\mathbf{k},\mu}e^{i\mathbf{k}\cdot\mathbf{x}} \tag{A1}$$

with

$$q_{\mathbf{k},\mu}{}^{*}=q_{-\mathbf{k},\mu}.$$

The $q_{\mathbf{k},\mu}$ are coordinates of the electromagnetic field, associated with oscillations of wave number, $\mathbf{k}$, and polarization direction, μ, where $\boldsymbol{\epsilon}_{\mathbf{k},\mu}$ is a unit vector normal to $\mathbf{k}$ and μ runs over two indices, corresponding to a pair of orthogonal directions of polarization. V is the volume of the box, which is assumed to be very large.

We also introduce the momenta $\prod_{\mathbf{k},\mu}=\partial q_{\mathbf{k},\mu}{}^{*}/\partial\dot{t}$, canonically conjugate[8] to the $q_{\mathbf{k},\mu}$. We have for the transverse part of the electric field

$$\mathfrak{E}(\mathbf{x})=-\frac{1}{c}\frac{\partial\mathbf{A}(\mathbf{x})}{\partial t}=-\left(\frac{4\pi}{Vc^2}\right)^{\frac{1}{2}}\sum_{\mathbf{k},\mu}\boldsymbol{\epsilon}_{\mathbf{k},\mu}\prod_{\mathbf{k},\mu}{}^{*}e^{i\mathbf{k}\cdot\mathbf{x}} \tag{A2}$$

and for the magnetic field

$$\mathfrak{H}(\mathbf{x})=\nabla\times\mathbf{A}=-(4\pi/V)^{\frac{1}{2}}i\sum_{\mathbf{k},\mu}(\mathbf{k}\times\boldsymbol{\epsilon}_{\mathbf{k},\mu})q_{\mathbf{k},\mu}e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{A3}$$

The Hamiltonian of the radiation field corresponds to a collection of independent harmonic oscillators, each with angular frequency, $\omega=kc$. This Hamiltonian is

$$H^{(R)}=\sum_{\mathbf{k},\mu}(\prod_{\mathbf{k},\mu}\prod_{\mathbf{k},\mu}{}^{*}+k^2c^2q_{\mathbf{k},\mu}q_{\mathbf{k},\mu}{}^{*}). \tag{A4}$$

Now, in our interpretation of the quantum theory, the quantity $q_{\mathbf{k},\mu}$ is assumed to refer to the actual value of the $\mathbf{k}$, μ Fourier component of the vector potential. As in the case of the electron, however, there is present an objectively real superfield that is a function of all the electromagnetic field coordinates $q_{\mathbf{k},\mu}$. Thus, we have

$$\Psi^{(R)}=\Psi^{(R)}(\cdots q_{\mathbf{k},\mu}\cdots). \tag{A5}$$

Writing $\Psi^{(R)}=R\exp(iS/\hbar)$, we obtain (in analogy with Paper I, Sec. 4)

$$\partial q_{\mathbf{k},\mu}/\partial t=\prod_{\mathbf{k},\mu}{}^{*}=\partial s/\partial q_{\mathbf{k},\mu}{}^{*}. \tag{A6}$$

The function $R(\cdots q_{\mathbf{k},\mu}\cdots)$ has two interpretations. First, it defines an additional quantum-mechanical term appearing in Maxwell's equations. To see the origin of this term, let us write the generalized Hamilton-Jacobi equation of the electromagnetic field, analogous to Paper I, Eq. (4),

$$\frac{\partial s}{\partial t}+\sum_{\mathbf{k},\mu}\frac{\partial s}{\partial q_{\mathbf{k},\mu}}\frac{\partial s}{\partial q_{\mathbf{k},\mu}{}^{*}}+\sum_{\mathbf{k},\mu}(kc)^2q_{\mathbf{k},\mu}q_{\mathbf{k},\mu}{}^{*}-\frac{\hbar^2}{2R}\sum_{\mathbf{k},\mu}\frac{\partial^2R(\cdots q_{\mathbf{k},\mu}\cdots)}{\partial q_{\mathbf{k},\mu}\partial q_{\mathbf{k},\mu}{}^{*}}=0. \tag{A7}$$

The equation of motion of $q_{\mathbf{k},\mu}$, derived from the Hamiltonian implied by Eq. (A7) becomes

$$\ddot{q}_{\mathbf{k},\mu}+k^2c^2q_{\mathbf{k},\mu}=\frac{\partial}{\partial q_{\mathbf{k},\mu}{}^{*}}\left(\frac{\hbar^2}{2R}\sum_{\mathbf{k}',\mu'}\frac{\partial^2R}{\partial q_{\mathbf{k}',\mu'}\partial q_{\mathbf{k}',\mu'}{}^{*}}\right). \tag{A8}$$

Since Maxwell's equations for empty space follow when the right-hand side is zero, we see that the "quantum-mechanical" terms can profoundly modify the behavior of the electromagnetic field. In fact, it is this modification which will contribute to the explana-

[8] See G. Wentzel, *Quantum Theory of Fields* (Interscience Publishers, Inc., New York, 1948).

tion of the ability of an oscillator, $q_{\mathbf{k},\mu}$, to transfer large quantities of energy and momentum rapidly even when $q_{\mathbf{k},\mu}$ is very small, for when $q_{\mathbf{k},\mu}$ is small, the right-hand side of Eq. (A8) may become very large.

The second interpretation of R is that as in Paper I, Eq. (5), it defines a conserved probability density that each of the $q_{\mathbf{k},\mu}$ has a certain specified value. From this fact, we see that although large transfers of energy and momentum to a radiation oscillator can occur in a short time when R is small, the probability of such a process is (as was also shown in Paper I, Sec. 7) very small.

In the lowest state (when no quanta are present) every oscillator is in the ground state. The super wave fields is then

$$\Psi_0^{(R)}=\exp[-\textstyle\sum_{\mathbf{k},\mu}(kcq_{\mathbf{k},\mu}q_{\mathbf{k},\mu}{}^*+\frac{1}{2}ikct)]. \quad \text{(A9)}$$

If the $\mathbf{k}'$, μ' oscillator is excited to the nth quantum state, the super wave field is

$$\Psi^{(R)}=h_n(q_{\mathbf{k}',\mu'})e^{-ink'ct}\Psi_0^{(R)}, \quad \text{(A10)}$$

where h_n is the nth hermite polynomial. As shown in Paper I, Sec. 5, the stationary states of such a system correspond to a quantized energy equal to the same value, $E_n=(n+\frac{1}{2})\hbar kc$, obtained from the usual interpretation. In nonstationary states, however, Eqs. (A7) and (A8) imply that the energy of each oscillator may fluctuate violently, as was also true of nonstationary states of the hydrogen atom (see Paper I, Sec. 7).

A nonstationary state of particular interest in the photoelectric and Compton effects is a state corresponding to the presence of an electromagnetic wave packet containing a single quantum. The super wave field for such a state is

$$\Psi_P{}^{(R)}=\textstyle\sum_{\mathbf{k},\mu}f_\mu(\mathbf{k}-\mathbf{k}_0)q_{\mathbf{k},\mu}e^{-ikct}\Psi_0^{(R)}, \quad \text{(A11)}$$

where $f_\mu(\mathbf{k}-\mathbf{k}_0)$ is a function that is large only near $\mathbf{k}=\mathbf{k}_0$ and the first hermite polynomial is represented by $q_{\mathbf{k},\mu}$, to which it is proportional.

To prove that Eq. (A11) represents an electromagnetic wave packet, we can evaluate the difference

$$\langle\Delta W\rangle_{Av}=\langle W\rangle_{Av}-\langle W_0\rangle_{Av}, \quad \text{(A12)}$$

where $\langle W(\mathbf{x})\rangle_{Av}$ is the actual mean energy density present (averaged over the ensemble), and $\langle W_0(\mathbf{x})\rangle_{Av}$ is the mean energy that would be present even in the ground state, because of zero-point fluctuations. We have

$$\langle W(\mathbf{x})\rangle_{Av}=\int\int\cdots\int\Psi_P{}^{*(R)}(\cdots q_{\mathbf{k},\mu}\cdots)\times\frac{[\mathfrak{E}^2(\mathbf{x})+\mathfrak{H}^2(\mathbf{x})]}{8\pi}\Psi_P{}^{(R)}(\cdots q_{\mathbf{k},\mu}\cdots)\times(\cdots dq_{\mathbf{k},\mu}\cdots). \quad \text{(A13)}$$

$$\langle W_0(\mathbf{x})\rangle_{Av}=\int\int\cdots\int\Psi_0^{*(R)}(\cdots q_{\mathbf{k},\mu}\cdots)\times\frac{[\mathfrak{E}^2(\mathbf{x})+\mathfrak{H}^2(\mathbf{x})]}{8\pi}\Psi_0^{(R)}(\cdots q_{\mathbf{k},\mu}\cdots)\times(\cdots dq_{\mathbf{k},\mu}\cdots). \quad \text{(A14)}$$

Obtaining $\mathfrak{E}(\mathbf{x})$ from Eq. (A2), $\mathfrak{H}(\mathbf{x})$ from Eq. (A3), $\Psi_P{}^{(R)}$ from Eq. (A10), $\Psi_0{}^{(R)}$ from Eq. (A9), we readily show that

$$\langle\Delta W(\mathbf{x})\rangle_{Av}=\textstyle\sum_{\mathbf{k},\mu}\sum_{\mathbf{k}',\mu'}f_\mu(\mathbf{k}-\mathbf{k}_0)f_{\mu'}(\mathbf{k}'-\mathbf{k}_0)\times e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}}\boldsymbol{\epsilon}_{\mathbf{k},\mu}\cdot\boldsymbol{\epsilon}_{\mathbf{k}',\mu'}. \quad \text{(A15)}$$

This means that the wave packet implies an excess over zero-point energy that is localized within a region in which the packet function, $\mathfrak{g}(\mathbf{x})$ is appreciable, where

$$\mathfrak{g}(\mathbf{x})=\textstyle\sum_{\mathbf{k},\mu}f_\mu(\mathbf{k}-\mathbf{k}_0)e^{i\mathbf{k}\cdot\mathbf{x}}\boldsymbol{\epsilon}_{\mathbf{k},\mu}. \quad \text{(A16)}$$

We are now ready to treat the photoelectric and Compton effects. The entire treatment is so similar to that of the Franck-Hertz experiment (Paper I, Sec. 7) that we need merely sketch it here. We begin by adding to the radiation Hamiltonian, $H^{(R)}$, the particle Hamiltonian,

$$H^{(P)}=(1/2m)[\mathbf{p}-(e/c)\mathbf{A}(\mathbf{x})]^2. \quad \text{(A17)}$$

(We restrict ourselves here to nonrelativistic treatment.) The photoelectric effect corresponds to the transition of a radiation oscillator from an excited state to the ground state, while the atomic electron is ejected, with an energy $E=h\nu-I$, where I is the ionization potential of the atom. The initial super wave field, corresponding to an incident packet containing only one quantum, plus an atom in the ground state is (see Eq. (A11))

$$\Psi_i=\psi_0(\mathbf{x})\exp(-iE_0t/\hbar)\Psi_0{}^{(R)}(\cdots q_{\mathbf{k},\mu}\cdots)\times\textstyle\sum_{\mathbf{k},\mu}f_\mu(\mathbf{k}-\mathbf{k}_0)q_{\mathbf{k},\mu}e^{-ikct}. \quad \text{(A18)}$$

By solving Schroedinger's equation for the combined system, we obtain an asymptotic wave field analogous to Paper I, Eq. (26), containing terms corresponding to the photoelectric effect. These terms, which must be added to Ψ_i, to yield the complete superfield, are (asymptotically)

$$\delta\Psi_a=\Psi_0{}^{(R)}(\cdots q_{\mathbf{k},\mu}\cdots)\sum_{\mathbf{k},\mu}f_\mu(\mathbf{k}-\mathbf{k}_0)\times\frac{\exp[i\mathbf{k}'\cdot\mathbf{r}-i\hbar(k'^2/2m)t]}{r}g_\mu(\theta,\phi,k'), \quad \text{(A19)}$$

where the energy of the outgoing electron is $E=\hbar^2k'^2/2m=\hbar kc+E_0$. The function $g_\mu(\theta,\phi,k')$ is the amplitude associated with the ψ-field of the outgoing electron. This quantity can be calculated from the matrix ele-

ment of the interaction term, $-(e/c)\mathbf{p}\cdot\mathbf{A}(\mathbf{x})$, by methods that are easily deducible from the usual perturbation theory.[6]

The outgoing electron packet has its center at $r=(\hbar k'/m)t$. Eventually, this packet will become completely separated from the initial electron wave function, $\psi_0(\mathbf{x})$. If the electron happens to enter the outgoing packet, the initial wave function can subsequently be ignored. The system then acts for all practical purposes as if its wave field were given by Eq. (A9), from which we conclude that the radiation field is in the ground state, while the electron has been liberated. It is readily shown that, as in the usual interpretation, the probability that the electron appears in the direction θ, ϕ can be calculated from $|g_\mu(\theta, \phi, k')|^2$ (see Paper I, Sec. 7).

To describe the Compton effect, we need only add to the super wave field the term corresponding to the appearance of an outgoing electromagnetic wave, as well as an outgoing electron. This part is asymptotically

$$\delta\Psi_b=\Psi_0^{(E)}(\cdots q_{\mathbf{k},\mu}\cdots)\sum_{\mathbf{k}',\mu'} f_\mu(\mathbf{k}-\mathbf{k}_0)$$

$$\times c_{\mathbf{k}',\mu'}{}^{\mathbf{k},\mu} q_{\mathbf{k}',\mu'} g_{\mathbf{k}',\mu'}{}^{\mathbf{k},\mu}(\theta,\phi)\frac{e^{ik''r}}{r}$$

$$\times\exp\left(-ik'ct-\frac{i\hbar k''^2 t}{2m}\right), \quad \text{(A20)}$$

where

$$(\hbar^2 k''^2/2m)+\hbar k'c=\hbar kc+E_0.$$

The quantity, $c_{\mathbf{k}',\mu'}{}^{\mathbf{k},\mu}$ is proportioned to the matrix element for a transition in which the $\mathbf{k}, \mu$-radiation oscillator falls from the first excited state, to the ground state, while the $\mathbf{k}', \mu'$-oscillator rises from the ground state to the first excited state. This matrix element is determined mainly by the term $(e^2/8mc^2)A^2(\mathbf{x})$ in the hamiltonian.

It is easily seen that the outgoing electron packet eventually becomes completely separated both from the initial wave field, $\Psi_s(\mathbf{x}, \cdots q_{\mathbf{k},\mu}\cdots)$, and from the packet for the photoelectric effect, $\delta\Psi_a$ [defined in Eq. (A19)]. If the electron should happen to enter this packet, then the others can be ignored, and the system acts for all practical purposes like an outgoing electron, plus an independent outgoing light quantum. The reader will readily verify that the probability that the light quantum $\mathbf{k}', \mu'$ appears along with an electron with angles θ, ϕ is precisely the same as in the usual interpretation.

APPENDIX B. A DISCUSSION OF INTERPRETATIONS OF THE QUANTUM THEORY PROPOSED BY DE BROGLIE AND ROSEN

After this article had been prepared, the author's attention was called to two papers in which an interpretation of the quantum theory similar to that suggested here was proposed, first by L. de Broglie,[9] and later by N. Rosen.[10] In both of these papers, it was suggested that if one writes $\psi=R\exp(is/\hbar)$, then one can regard R^2 as a probability density of particles having a velocity, $\mathbf{v}=\nabla s/m$. De Broglie regarded the ψ-field as an agent "guiding" the particle, and therefore referred to ψ as a "pilot wave." Both of these authors came to the conclusion that this interpretation could not consistently be carried through in those cases in which the field contained a linear combination of stationary state wave functions. As we shall see in this appendix, however, the difficulties encountered by the above authors could have been overcome by them, if only they had carried their ideas to a logical conclusion.

De Broglie's suggestions met strong objections on the part of Pauli,[11] in connection with the problem of inelastic scattering of a particle by a rigid rotator. Since this problem is conceptually equivalent to that of inelastic scattering of a particle by a hydrogen atom, which we have already treated in Paper I, Sec. 7, we shall discuss the objections raised by Pauli in terms of the latter example.

Now, according to Pauli's argument, the initial wave function in the scattering problem should be $\Psi=\exp(i\mathbf{p}_0\cdot\mathbf{y}/\hbar)\psi_0(\mathbf{x})$. This corresponds to a stationary state for the combined system, in which the particle momentum is $\mathbf{p}_0$, while the hydrogen atom is in its ground state, with a wave function, $\psi_0(\mathbf{x})$. After interaction between the incident particle and the hydrogen atom, the combined wave function can be represented as

$$\Psi=\sum_n f_n(\mathbf{y})\psi_n(\mathbf{x}), \quad \text{(B1)}$$

where $\psi_n(\mathbf{x})$ is the wave function for the nth excited state of the hydrogen atom, and $f_n(\mathbf{y})$ is the associated expansion coefficient. It is easily shown[12] that asymptotically, $f_n(\mathbf{y})$ takes the form of an outgoing wave, $f_n(\mathbf{y})\sim g_n(\theta,\phi)e^{ik_n r}/r$, where $(\hbar k_n)^2/2m=[(\hbar k_0)^2/2m]+E_n-E_0$. Now, if we write $\psi=R\exp(iS/\hbar)$, we find that the particle momenta, $p_x=\nabla_x S(\mathbf{x},\mathbf{y})$ and $\mathbf{p}_y=\nabla_y S(\mathbf{x},\mathbf{y})$, fluctuate violently in a way that depends strongly on the position of each particle. Thus, neither atom nor the outgoing particle ever seem to approach a stationary energy. On the other hand, we know from experiment that both the atom and the outgoing particle do eventually obtain definite (but presumably unpredictable) energy values. Pauli therefore concluded that the interpretation proposed by de Broglie was untenable. De Broglie seems to have agreed with the conclusion, since he subsequently gave up his suggested interpretation.[9]

---

[9] L. de Broglie, *An Introduction to the Study of Wave Mechanics* (E. P. Dutton and Company, Inc., New York, 1930), see Chapters 6, 9, and 10.

[10] N. Rosen, J. Elisha Mitchel Sci. Soc. **61**, Nos. 1 and 2 (August, 1945).

[11] *Reports on the 1927 Solvay Congress* (Gauthiers-Villars et Cie., Paris, 1928), see p. 280.

[12] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions* (Clarendon Press, Oxford, 1933).

Our answer to Pauli's objection is already contained in Paper I, Sec. 7, as well as in Sec. 2 of this paper. For as is well known, the use of an incident plane wave of infinite extent is an excessive abstraction, not realizable in practice. Actually, both the incident and outgoing parts of the ψ-field will always take the form of bounded packets. Moreover, as shown in Paper I, Sec. 7, all packets corresponding to different values of n will ultimately obtain classically describable separations. The outgoing particle must enter one of these packets, and it will remain with that particular packet thereafter, leaving the hydrogen atom in a definite but correlated stationary state. Thus, Pauli's objection is seen to be based on the use of the excessively abstract model of an infinite plane wave.

Although the above constitutes a complete answer to Pauli's specific objections to our suggested interpretation, we wish here to amplify our discussion somewhat, in order to anticipate certain additional objections that might be made along similar lines. For at this point, one might argue that even though the wave packet is bounded, it can nevertheless in principle be made arbitrarily large in extent by means of a suitable adjustment of initial conditions. Our interpretation predicts that in the region in which incident and outgoing ψ-waves overlap, the momentum of each particle will fluctuate violently, as a result of corresponding fluctuations in the "quantum-mechanical" potential produced by the ψ-field. The question arises, however, as to whether such fluctuations can really be in accord with experimental fact, especially since in principle they could occur when the particles were separated by distances much greater than that over which the "classical" interaction potential, $V(\mathbf{x}, \mathbf{y})$, was appreciable.

To show that these fluctuations are not in disagreement with any experimental facts now available, we first point out that even in the usual interpretation the energy of each particle cannot correctly be regarded as definite under the conditions which are assumed here, namely, that the incident and outgoing wave packets overlap. For as long as interference between two stationary state wave function is possible, the system acts as if it, in some sense, covered both states simultaneously.[13] In such a situation, the usual interpretation implies that a precisely defined value for the energy of either particle is meaningless. From such a wave function, one can predict only the probability that if the energy is measured, a definite value will be obtained. On the other hand, the very experimental conditions needed for measuring the energy play a key role in making a definite value of the energy possible because the effect of the measuring apparatus is to destroy interference between parts of the wave function corresponding to different values of the energy.[14]

[13] Reference 2, Chapter 16, Sec. 25.
[14] Reference 2, Chapter 6, Secs. 3 to 8; Chapter 22, Secs. 8 to 10.

In our interpretation, the overlap of incident and outgoing wave packets signifies not that the precise value of the energy of either particle can be given no meaning, but rather that this value fluctuates violently in an, in practice, unpredictable and uncontrollable way. When the energy of either particle is measured, however, then our interpretation predicts, in agreement with the usual interpretation, that the energy of each particle will become definite and constant, as a result of the effects of the energy-measuring apparatus on the observed system. To show how this happens, let us suppose that the energy of the hydrogen atom is measured by means of an interaction in which the "classical" potential, V, is a function only of the variables associated with the electron in the hydrogen atom and with the apparatus, but is not a function of variables associated with the outgoing particle. Let z be the coordinate of the measuring apparatus. Then as shown in Sec. 2, interaction with an apparatus that measures the energy of the hydrogen atom will transform the Ψ-function (B1), into

$$\Psi=\sum_n f_n(\mathbf{y})\psi_n(\mathbf{x})g_0(z-aEnt/\hbar^2). \qquad \text{(B2)}$$

Now, we have seen that if the product at is large enough to make a distinct measurement possible, packets corresponding to different values of n will ultimately obtain classically describable separations in z space. The apparatus variable, z, must enter one of these packets; and, thereafter, all other packets can for practical purposes be ignored. The hydrogen atom is then left in a state having a definite and constant energy, while the outgoing particle has a correspondingly definite but correlated constant value for its energy. Thus, we find that as with the usual interpretation, our interpretation predicts that whenever we measure the energy of either particle by methods that are now available, a definite and constant value will always be obtained. Nevertheless, under conditions in which incident and outgoing wave packets overlap, and in which neither particle interacts with an energy-measuring device, our interpretation states unambiguously that real fluctuations in the energy of each particle will occur. These fluctuations are moreover, at least in principle, observable (for example, by methods discussed in Sec. 6). Meanwhile, under conditions in which we are limited by present methods of observation, our interpretation leads to predictions that are precisely the same as those obtained from the usual interpretation, so that no experiments supporting the usual interpretation can possibly contradict our interpretation.

In his book,[9] de Broglie raises objections to his own suggested interpretation of the quantum theory, which are very similar to those raised by Pauli. It is therefore not necessary to answer de Broglie's objections in detail here, since the answer is essentially the same as that which has been given to Pauli. We wish, however,

to add one point. De Broglie assumes that not only electrons, but also light quanta, are associated with particles. A consistent application of the interpretation suggested here requires, however, as shown in Appendix A, that light quanta be described as electromagnetic wave packets. The only precisely definable quantities in such a packet are the Fourier components, $q_{k,\mu}$, of the vector potential and the corresponding canonically conjugate momenta, $\Pi_{k,\mu}$. Such packets have many particle-like properties, including the ability to transfer rapidly a full quantum of energy at great distances. Nevertheless, it would not be consistent to assume the existence of a "photon" particle, associated with each light quantum.

We shall now discuss Rosen's paper briefly.[10] Rosen gave up his suggested interpretation of the quantum theory, because of difficulties arising in connection with the interpretation of standing waves. In the case of the stationary states of a free particle in a box, which we have already discussed in Sec. 8, our interpretation leads to the conclusion that the particle is standing still. Rosen did not wish to accept this conclusion, because it seemed to disagree with the statement of the usual interpretation that in such a state the electron is moving with equal probability that the motion is in either direction. To answer Rosen's objections, we need merely point out again that the usual interpretation can give no meaning to the motion of particles in a stationary state; at best, it can only predict the probability that a given result will be obtained, if the velocity is measured. As we saw in Sec. 8, however, our interpretation leads to precisely the same predictions as are obtained from the usual interpretation, for any process which could actually provide us with a measurement of the velocity of the electron. One must remember, however, that the value of the momentum "observable" as it is now "measured" is not necessarily equal to the particle momentum existing before interaction with the measuring apparatus took place.

We conclude that the objections raised by Pauli, de Broglie, and Rosen, to interpretations of the quantum theory similar to that suggested here, can all be answered by carrying every aspect of our suggested interpretation to its logical conclusion.

PHYSICAL REVIEW VOLUME 85, NUMBER 2 JANUARY 15, 1952

Neutrons from the Disintegration of the Separated Isotopes of Silicon by Deuterons*

C. E. MANDEVILLE, C. P. SWANN, AND S. D. CHATTERJEE†
Bartol Research Foundation of the Franklin Institute, Swarthmore, Pennsylvania

AND

D. M. VAN PATTER‡
Department of Physics and Laboratory for Nuclear Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received September 14, 1951)

The neutron spectra from deuterons on silicon have been studied by the method of recoil protons and photographic plates. Thick isotopic targets of the three separated isotopes of silicon were irradiated by deuterons of energy 1.4 Mev, supplied by the Bartol Van de Graaff statitron, observations being carried out at angles of zero and ninety degrees with the incident deuterons. Q-values, from which energy levels in the residual nuclei of phosphorous may be calculated, are as follows:

Reaction	Q-values (Mev)
$Si^{28}(d,n)P^{29}$	0.29
$Si^{29}(d,n)P^{30}$	3.27, 2.52, 1.81, 1.27
$Si^{30}(d,n)P^{31}$	4.92, 4.59, 3.73, 2.70, 1.51

The estimated probable error in the Q-values is 40 kev.

THERE are three stable isotopes of silicon, Si^{28}, Si^{29}, and Si^{30}, having respectively relative abundances of 92.28 percent, 4.67 percent, and 3.05 percent. When these elements are irradiated by deuterons, neutrons are emitted in the following three reactions.

(1) $Si^{28}+d\rightarrow P^{29}+n^1+Q_1$,

(2) $Si^{29}+d\rightarrow P^{30}+n^1+Q_2$,

(3) $Si^{30}+d\rightarrow P^{31}+n^1+Q_3$.

Naturally occurring silicon has been previously bombarded with deuterons to observe the neutron spectra.[1] However, because of the element of ambiguity introduced by the presence of the mixture of isotopes, the data are difficult of interpretation. To reinvestigate the above three reactions, quantities of the separated isotopes, in the form of silicon dioxide, were obtained from the Y-12 plant, Carbide and Carbon Chemicals Division, Union Carbide and Carbon Corporation, Oak Ridge, Tennessee. The mass analyses of the various targets are shown in Table I.

* Assisted by the joint program of the ONR and AEC.

† Calcutta, India. Guest physicist, Bartol Research Foundation, 1951.

‡ The contribution of D. M. Van Patter to this article consists of the preparation of the Appendix.

[1] R. A. Peck, Jr., Phys. Rev. 73, 947 (1948).

TABLE I. Isotopic analysis of target materials.

Si^{28}		Si^{29}		Si^{30}	
Isotope	Percent	Isotope	Percent	Isotope	Percent
28	99.4	28	29.1	28	40.97
29	0.4	29	68.6	29	2.44
30	0.2	30	2.2	30	56.59

Thick isotopic targets of the separated isotopes of silicon were irradiated by deuterons of energy 1.40 Mev, supplied by the Bartol Van de Graaff statitron. The neutrons emitted in the three reactions above were recorded in Eastman NTA plates located at zero and ninety degrees with the incident deuterons.

In evaluating the energies of the neutron groups of all of the three reactions above, two procedures have been employed. One has been to apply the range-energy relation of Lattes, Fowler, and Cuer[2] to NTA plates, increasing the observed energy by two percent to allow for an acceptance angle of twelve degrees in the forward direction; the other has consisted in decreasing the result of the first method by an amount indicated by a recently obtained calibration curve[3] for NTA plates, and increasing it to take into account the use of a thick target.[4] The latter procedure has been considered to give the preferable result. The procedures of evaluation described above will be referred to respectively as Method I and Method II.

$Si^{28}(d,n)P^{29}$

The energy spectrum of the acceptable recoil protons of the neutrons emitted in the reaction $Si^{28}(d,n)P^{29}$ is shown in Fig. 1. Observations were carried out at zero degrees and ninety degrees in the laboratory system of coordinate axes. As indicated in the figure, two groups of neutrons appear at both angles of observation. The group of lower energy is assigned to the reaction $C^{12}(d,n)N^{13}$. The more energetic is thought to be related to the formation of P^{29} in its ground state. In the forward direction, high energy neutrons of low intensity

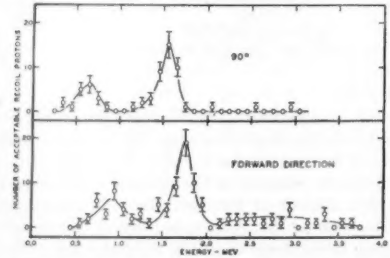


FIG. 1. Recoil protons knocked on in the forward direction by the neutrons from $Si^{28}(d,n)P^{29}$.

TABLE II. Ground-state Q-values for $Si^{28}(d,n)P^{29}$.

	Method I	Method II
Q-value (Mev)	0.36 ± 0.04	0.29 ± 0.04

are present. These are thought to result from the bombardment of "low Z" contaminants on the walls of the vacuum tube and of the magnet box between the poles of the beam focusing magnet of the Van de Graaff statitron. Plates in the forward direction are sensitive to neutrons emanating from these points, whereas the plates at ninety degrees to the beam are perpendicular to their direction of emission and hence do not record their recoil protons in the acceptable sense. This point is born out by the fact that no appreciable evidence of the presence of neutrons beyond those of the ground state is apparent in the data obtained at ninety degrees. The data of Fig. 1 are summarized in Table II.

The Q-values of Table II are means of those calculated at each of the two angles of observation. It is to be noted that they are in disagreement with the previously reported Q-value of -0.80 ± 0.10 Mev,[1] and with those values calculated from other reactions and from mass values (see Appendix).

$Si^{29}(d,n)P^{30}$

The spectra of acceptable recoil protons of the neutrons emanating from $Si^{29}(d,n)P^{30}$ are shown in Fig. 2. The most energetic group[5] is assigned to $N^{14}(d,n)O^{15}$,

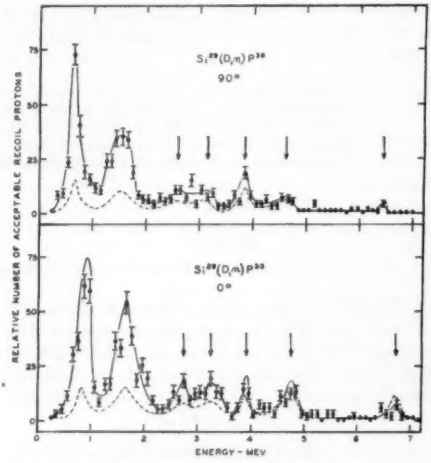


FIG. 2. Energy spectrum of the acceptable recoil protons of the neutrons from the reaction $Si^{29}(d,n)P^{30}$.

[2] Lattes, Fowler, and Cuer, Proc. Phys. Soc. (London) **59**, 883 (1947).

[3] Richards, Johnson, Ajzenberg, and Laubenstein, Phys. Rev. **83**, 994 (1951).

[4] S. C. Snowdon (to be published).

[5] The neutron group of the nitrogen reaction appears at zero degrees and ninety degrees in the data relating to $Si^{29}(d,n)P^{30}$ and in the forward direction alone in measurements on $Si^{30}(d,n)P^{31}$. The absence of this group at ninety degrees in the latter reaction

the least energetic one to $C^{12}(d,n)N^{13}$. The ground-state group of the reaction $Si^{28}(d,n)P^{29}$, the reaction of the preceding discussion, is also clearly present. The intermediate groups are assigned to the reaction $Si^{29}(d,n)P^{30}$. The data of Fig. 2 are summarized in Table III.

The energy levels of P^{30} in Table III are calculated from the preferred values of Method II.

$Si^{30}(d,n)P^{31}$

The acceptable recoil proton groups of the neutrons from $Si^{30}(d,n)P^{31}$ are shown in Fig. 3.[6] Here again, the most energetic group, appearing at 6.75 Mev in the forward direction, is assigned to $N^{14}(d,n)O^{15}$; the least energetic one to $C^{12}(d,n)N^{13}$. Again, the ground-state group of $Si^{28}(d,n)P^{29}$ is present. The remaining groups are assigned to $Si^{30}(d,n)P^{31}$. The data of Fig. 3 are summarized in Table IV.§

The nuclear energy levels of P^{31} have been calculated from the preferred Q-values of Method II.

TABLE III. Q-Values of $Si^{29}(d,n)P^{30}$ and energy levels of P^{30}.

Q-values (Mev) Method I	Q-values (Mev) Method II	Excitation levels of P^{30}
1.41±0.04	1.27±0.04	2.00±0.06
1.97±0.04	1.81±0.04	1.46±0.06
2.68±0.04	2.52±0.04	0.75±0.06
3.51±0.04	3.27±0.04	0.00

CONCLUSIONS

The energies of the neutron groups from deuterons on the separated isotopes of silicon have been measured. In many instances the calculated energy levels of the residual nuclei of phosphorous agree with those already obtained (see reference 1, Table II). The bearing of the measurements upon the masses and reaction energies of the silicon region are discussed in the appendix.

The writers wish to acknowledge the continued interest of Dr. W. F. G. Swann, Director of the Bartol Research Foundation.

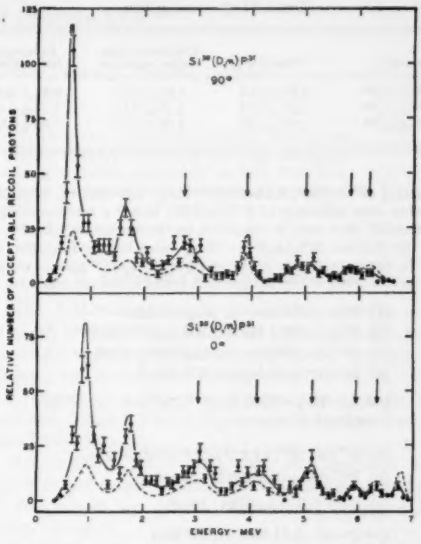


FIG. 3. Recoil protons of the neutrons emitted in the reaction $Si^{30}(d,n)P^{31}$.

APPENDIX: DISCUSSION OF RESULTS

Recent precise measurements[7,8] of nuclear Q-values involving nuclei with $Z=10$ to 16 have led to a revision of masses in that region.[8,10] However, in many cases, the most accurate method of estimating a given reaction energy involves the use of mass differences rather than a table of masses. It is of interest to compare the Q-values of the ground states measured by Mandeville, Swann, and Chatterjee with the reaction energies which can be predicted from the measurements of other workers.

The Reaction $Si^{28}(d,n)P^{29}$.—The end point of the beta-transition $P^{29}\xrightarrow{\beta^+}Si^{29}$ has been measured as 3.63±0.07 Mev[11] by means of cloud chamber and magnetic field. Except for the results of the present paper relating to reaction $Si^{28}(d,n)P^{29}$, the beta-transition above is the only link connecting the nucleus of P^{29} with other nuclei. Combining the total disintegration energy of the beta transition, 4.65±0.07 Mev, with the reaction energy of the

TABLE IV. Q-values of $Si^{30}(d,n)P^{31}$ and energy levels of P^{31}.

Q-values (Mev) Method I	Q-values (Mev) Method II	Excitation levels of P^{31} (Mev)
1.64±0.04	1.51±0.04	3.41±0.06
2.28±0.04	2.70±0.04	2.22±0.06
3.95±0.04	3.73±0.04	1.19±0.06
4.79±0.04	4.59±0.04	0.33±0.06
5.16±0.04	4.92±0.04	0.00

is not considered significant, because statistical difficulties are encountered in accumulating acceptable long tracks. From the several spectra, an average Q-value of 5.11±0.04 Mev is calculated for $N^{14}(d,n)O^{15}$, using Method II of the text. This result is in good agreement with the previously reported energy release of 5.15±0.05 Mev [W. M. Gibson and D. L. Livesey, Proc. Phys. Soc. (London) **60**, 523 (1948)].

[6] The broken line of Figs. 2 and 3 represents the actual distribution of neutron energies corrected for variation with energy of the n-p scattering cross section and acceptance probability.

§ *Note added in proof:* In each of the three figures, the data at zero and ninety degrees have not been normalized to take into account plate area scanned. No information is thus presented concerning the angular distributions of the neutron groups. The fluctuation of the intensity of the carbon group relative to those of the other neutron groups is thought to be related to statitron performance. Excessive sparking and breakdown in the vacuum tube appear to increase the carbon contamination.

[7] Strait, Van Patter, Buechner, and Sperduto, Phys. Rev. **81**, 747 (1951).

[8] W. W. Buechner, *et al.*, M.I.T. report, Laboratory for Nuclear Science and Engineering (May 31, 1951).

[9] H. T. Motz, Phys. Rev. **81**, 1061 (1951).

[10] A. H. Wapstra, to be published.

[11] White, Creutz, Delsasso, and Wilson, Phys. Rev. **59**, 63 (1941).

TABLE AI. Q-values in Mev.

Reaction	Observed	Calculated from nuclear reactions	Calculated from masses
$Si^{28}(d,n)P^{29}$	0.29 ± 0.04	0.81 ± 0.07	0.81 ± 0.09
$Si^{29}(d,n)P^{30}$	3.27 ± 0.04	3.25 ± 0.18	3.24 ± 0.11
$Si^{30}(d,n)P^{31}$	4.92 ± 0.04	5.08 ± 0.02	5.03 ± 0.06

reaction $Si^{28}(d,p)Si^{29}$, 6.246 ± 0.007 Mev,[7] and with the neutron-proton mass difference of 0.782 ± 0.001 Mev,[12] a reaction energy 0.81 ± 0.07 Mev may be calculated for the reaction $Si^{28}(d,n)P^{29}$.

The Reaction $Si^{29}(d,n)P^{30}$.—The Q-value for the formation of P^{30} in the ground state in the reaction $Si^{29}(d,n)P^{30}$ may be calculated from each of two independent cycles of nuclear reactions.

I. (a) $P^{31}+\gamma\rightarrow P^{30}+n-(12.37\pm0.2)$ Mev[13,14]
(b) $P^{31}+p\rightarrow Si^{28}+He^4+(1.909\pm0.015)$ Mev[15]
(c) $Si^{28}+d\rightarrow Si^{29}+p+(6.246\pm0.007)$ Mev[7]
(d) $2d\rightarrow He^4+(23.834\pm0.007)$ Mev[12]

$Q=(2d-He^4)-1.909$ Mev -6.246 Mev -12.37 Mev
$=3.31\pm0.20$ Mev

II. (a) $Si^{29}+d\rightarrow Si^{30}+p+(8.388\pm0.013)$ Mev[8]
(b) $P^{30}\xrightarrow{\beta^+}Si^{30}+(4.52\pm0.35)$ Mev[16]
(c) $n\xrightarrow{\beta^-}p+(0.782\pm0.001)$ Mev[12]

$Q=(p-n)-4.52$ Mev $+8.388$ Mev
$=3.09\pm0.35$ Mev

A weighted average of results I and II is 3.25 ± 0.18 Mev.

The Reaction $Si^{30}(d,n)P^{31}$.—In the case of the reaction $Si^{30}(d,n)P^{31}$, precise estimates of the energy release can be made by two independent methods.

I. (a) $Si^{30}+d\rightarrow Si^{31}+p+(4.367\pm0.010)$ Mev[7]
(b) $Si^{31}\xrightarrow{\beta^-}P^{31}+(1.51\pm0.01)$ Mev[17]
(c) $n\xrightarrow{\beta^-}p+(0.782\pm0.001)$ Mev[12]

$Q=(p-n)+4.367$ Mev $+1.51$ Mev
$=5.095\pm0.015$ Mev

II. (a) $Si^{29}+d\rightarrow Si^{30}+p+(8.388\pm0.013)$ Mev[8]
(b) $Si^{28}+d\rightarrow Si^{29}+p+(6.246\pm0.007)$ Mev[7]
(c) $P^{31}+p\rightarrow Si^{28}+He^4+(1.909\pm0.015)$ Mev[15]
(d) $d\rightarrow p+n-(2.225\pm0.002)$ Mev[12]
(e) $2d\rightarrow He^4+(23.834\pm0.007)$ Mev[12]

$Q=(2d-He^4)+(d-n-p)-6.246$ Mev
-8.388 Mev -1.909 Mev
$=5.066\pm0.022$ Mev

An arithmetical average of the results of I and II is 5.08 ± 0.02 Mev.

The foregoing calculated values are listed in Table AI of this Appendix together with values estimated from a table of binding energies recently computed by Wapstra.[10] The agreement between these values is not surprising, because Wapstra's binding energies are based to a large extent upon nuclear Q-values.

It may be seen from this table that the observed Q-value for the reaction $Si^{29}(d,n)P^{30}$ is in agreement with other measurements, and, in fact, establishes the mass of P^{30} with greater precision than was previously possible. The observed energy release, 4.92 ± 0.04 Mev, for the reaction $Si^{30}(d,n)P^{31}$ is somewhat lower than the more accurate predicted value of 5.08 ± 0.02 Mev; however, it seems probable that the ground-state group of the reaction has been observed. In the case of $Si^{28}(d,n)P^{29}$, the observed energy release of 0.29 ± 0.04 Mev is in distinct disagreement with the estimated value of 0.81 ± 0.07 Mev. This discrepancy can be attributed to either an inaccuracy in the early measurement[11] of the end point of the transition $P^{29}\xrightarrow{\beta^+}Si^{29}$ or to the fact that the ground state of the reaction $Si^{28}(d,n)P^{29}$ has not been observed in the measurements of the present paper.

[10] Li, Whaling, Fowler, and Lauritsen, Phys. Rev. **83**, 512 (1951).
[11] McElhinney, Hanson, Becker, Duffield, and Diven, Phys. Rev. **75**, 542 (1949).
[12] A. Katz and L. Penfold, Phys. Rev. **81**, 815 (1951).
[13] D. M. Van Patter and P. M. Endt, preliminary value.
[14] C. Magnan, Ann. phys. **15**, 5 (1941).
[15] H. W. Newson, Phys. Rev. **51**, 624 (1937).

Radiative Capture of Protons by C^{13}*

JOHN D. SEAGRAVE†
Kellogg Radiation Laboratory, California Institute of Technology, Pasadena, California
(Received October 9, 1951)

The reaction $C^{13}(p,\gamma)N^{14}$ has been studied experimentally for proton bombarding energies between 0.4 and 2.7 Mev. Targets of carbon enriched in the isotope C^{13} were prepared, and thick and thin target excitation functions obtained. Five resonances were found, corresponding to excited states of N^{14} at 8.05, 8.62, 8.70, 9.18, and 9.49 Mev. The radiation spectrum from these levels was found to involve several branches of decay and transitions cascading through intermediate levels.

I. INTRODUCTION

THERE exists an important class of proton-induced reactions in light nuclei in which radiative capture is the only transmutation of the target nucleus energetically possible for proton energies up to several Mev. Three such (p,γ) reactions figure in the "carbon cycle" of energy production in the sun and hotter stars. Until quite recently, investigation of these reactions had been limited to proton energies between 0.2 and 1 Mev. Hornyak *et al.* give bibliography (to June, 1950) and energy level diagrams for these nuclei in their review article.[1] The nuclei N^{13} and O^{15} are positron emitters, and recent investigations[2,3] of proton capture by C^{12} and N^{14} have taken advantage of the β^+-activity produced in the targets to measure the small (p,γ) yield. In the case of the reaction $C^{13}(p,\gamma)N^{14}$, which is the principal subject of this paper, the residual nucleus is stable, and it is necessary to measure the gamma-radiation directly.

Unpublished work by Fowler and Lauritsen at this laboratory with a thick sample of C^{13}-enriched lampblack was continued above 1.3 Mev on the new electrostatic accelerator by Day and Perry, who discovered a very sharp resonance in the vicinity of 1.76 Mev, and a continuous increase in yield extending from the known resonance at 0.55 Mev. It was recognized that the sharp resonance should appear superimposed on the broader resonance at 1.7 Mev due to C^{12}, even in a target of normal carbon. This was found to be the case by Day, using a thin lampblack target. He was in fact able to obtain angular distributions for both resonances separately.[4]

The present work has involved preparation of thick and thin targets highly enriched in C^{13}, and an investigation of the energy levels and γ spectrum of N^{14} for proton energies from 0.4 to 2.7 Mev.

II. EXPERIMENTAL TECHNIQUES

The 3-Mev electrostatic accelerator recently constructed at the Kellogg Radiation Laboratory provided a steady source of protons, continuously variable in energy, and maintained homogeneous to better than 0.1 percent by a 90° magnetic analyzer of a double focusing design.[5] This equipment will be described in detail elsewhere.[6]

An auxiliary pump and cold trap between the analyzer and target chamber minimized contamination of the targets by normal carbon and oxygen from pump oil vapor.

The basic detection unit was a system of three Geiger tubes[7] of 30 mg/cm² glass wall and a sensitive region ¾ in. in diameter×3½ in. long. These were placed in a geometrical arrangement standardized at this laboratory.[8] Secondary electrons produced in a thick aluminum converter which pass through both the front counter and either one of the two rear counters are registered by standard coincidence circuits as coincidence counts. The spectrum of the radiation may be inferred by studying the reduction in coincidence rate as a function of the thickness of aluminum absorber placed directly in front of the rear counters. For the geometry employed, the ratio of coincidence counts to front counts with no absorber is 0.30±0.03 for 3- to 9-Mev gamma-radiation. The counter system was enclosed in a lead shield, and receptacle for a standard ThC″ γ-source permitted periodic checks on counter operation.

The thick target was simply prepared by compressing C^{13}-enriched lampblack into a ⅜-in. recess milled in a silver blank. It was found necessary to wet the lampblack with a dilute solution of shellac in alcohol. After evaporation of the solvent, the target material could be pressed into a pellet within the recess.

It was clear that thin lampblack targets were quite nonuniform and gave rise to nonuniformities in proton energy loss, and in addition could not be made thin enough for convenient investigation of narrow resonances. Preparation of very thin targets of high uni-

* This work was assisted by the joint program of the ONR and AEC.
† Now at Los Alamos Scientific Laboratory, Los Alamos, New Mexico.
[1] Hornyak, Lauritsen, Morrison, and Fowler, Revs. Modern Phys. **22**, 291 (1950).
[2] J. D. Seagrave, Phys. Rev. **84**, 1219 (1951).
[3] D. B. Duncan and J. E. Perry, Phys. Rev. **82**, 809 (1951).
[4] R. B. Day and J. E. Perry, Phys. Rev. **81**, 662(A) (1951).
[5] D. B. Duncan, Phys. Rev. **76**, 587(A) (1949).
[6] J. E. Perry, Rev. Sci. Instr. (to be published).
[7] Made by Radiation Counter Laboratories.
[8] Fowler, Lauritsen, and Lauritsen, Revs. Modern Phys. **20**, 236 (1948).

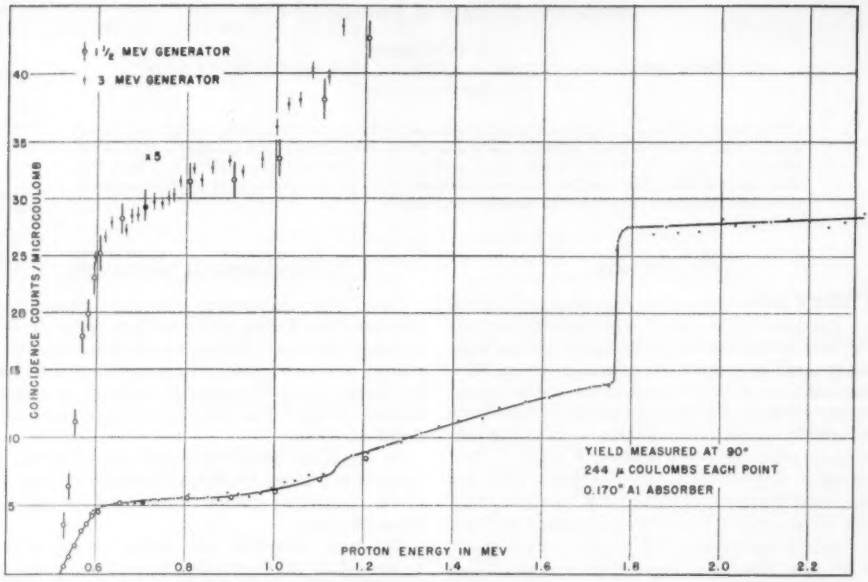


FIG. 1. Thick target yield of high energy radiation from $C^{13}(p,\gamma)N^{14}$ measured at 90° to the axis of the proton beam.

formity was achieved by "cracking" a hydrocarbon vapor or gas on a heated metal surface, leaving an adherent film of carbon. Phillips and Richardson have reported[9] two techniques, using a high temperature oven for cracking benzene on silver foils and later an induction heater to crack C^{13}-enriched CH_3I on nickel disks. The targets used at this laboratory were prepared simply by passing current through a thin tantalum strip suspended in a chamber filled to about $\frac{1}{10}$ atmosphere with CH_4, or (normal or enriched) CH_3I. Cracking proceeds satisfactorily at a temperature corresponding to a bright red color of the tantalum. On the polished Ta surface a thin film was quite invisible, but was detected and measured by the gamma-yield at the 1.76-Mev proton resonance as a result of C^{13}. By comparison with lampblack and graphite targets, these films were shown to be nearly pure carbon. The 1.76-Mev resonance was very convenient to use for the examination of the targets, since the coincidence counting rate from the 3.5-Mev radiation from N^{13} could be attenuated markedly compared to the 9.2-Mev radiation from N^{14} by the use of an appropriate absorber.

Trouble with "blistering" of the film was greatly reduced by the use of a sliding vacuum-fitting in the cracking chamber which maintained the Ta strips in tension during heating, to prevent buckling. It was also necessary to use vapor pressures less than 10 cm Hg, and to reduce temperature slowly. However, by encouraging blistering, it was found possible to remove sections intact as large as 5×5 mm, which should be of interest when an unsupported target is desired. In appearance these "foils" greatly resemble thin beryllium foils. The irregularities of thickness observed with the first films made were greatly reduced by careful polishing and cleaning of the tantalum. It was found unnecessary to outgas the Ta in a high vacuum, and it was merely exposed to a glow discharge while the chamber was being evacuated with the forepump alone (to about 5 microns). Although the voluminous production of I_2 vapor was a nuisance in the cracking chamber, it served as an index of the cracking process and targets made from normal CH_3I showed no detectable difference from those made from methane. Methyl iodide enriched to an absolute concentration of 61 atom percent C^{13} was obtained,[10] and a series of thin targets made from both normal and enriched methyl iodide. The three most satisfactory "C^{13}" targets were 16, 8, and <1 kev thick.

III. THE EXCITATION FUNCTIONS

The excitation function obtained with the C^{13}-enriched lampblack target is shown in Fig. 1. Since N^{14} produced is excited to about 9 Mev while N^{12} is excited

[9] G. C. Phillips and J. E. Richardson, Rev. Sci. Instr. **21**, 885 (1950).

[10] From Distillation Products Industries, Eastman Kodak Company. With conservative handling, a large number of targets can be made from 0.1 g of C^{13} in this form.

to about 3 Mev, an 0.150-inch aluminum absorber[11] was employed in the coincidence set-up described previously. This thickness of absorber reduced coincidence counts from N^{13} radiation to a negligible level, while attenuating those from 9-Mev radiation only about 50 percent. The coincidence excitation function shown is thus a measure of the high energy yield only. The bombardment was carried down to 600 kev, as low as it was convenient to operate the accelerator. The low energy data was in agreement with the data obtained by Fowler and Lauritsen on the 1½-Mev accelerator for proton energies of 400 to 1200 kev. The two sets of data were adjusted in scale to match at 700 kev, as shown.

The strong sharp resonance at 1.76 Mev manifests itself as an abrupt "step." The "rise above resonance" observed in the low energy data is seen to exhibit the general features of an extremely broad resonance centered near 1.3 Mev, and extending over several Mev. A small bump appears at 1.16 Mev, and is more marked in the front counter data (not shown), as is a similar bump at 2.1 Mev which is not resolved in the coincidence curve at all. These weak resonances thus give much softer radiation than do the principal resonances.

With the original thick target, Day and Perry determined the location of the sharp resonance as 1754±3 kev and its width as 2.5±0.5 kev. A later value of 1757±3 kev on the same target may have been shifted up by a layer of contamination deposited by the beam. It was hoped to obtain more precise values with the more highly enriched and uniform thin targets, but difficulties with regulation of the magnetic analyzer arose and made it difficult to validate a more precise determination of the energy. A sharper value for the width will be discussed below. It is planned to carry out this important experiment as soon as the electrostatic analyzer now being tested can be calibrated. Indeed, the resonance should prove very useful as a standard itself, as it is believed to be the sharpest proton resonance now known in the light nuclei ($Z \leq 10$), and its high energy radiation makes detection convenient. We shall continue to refer to it here as the "1.76-Mev resonance."

Sections of several C^{13}-coated tantalum strips were selected and their "profiles" examined by means of the sharp resonance. A target 16 kev thick at 1.76 Mev which was particularly uniform was used in most of the following work.

Preliminary thin target data confirmed the existence of resonances at 1.16 and 2.10 Mev and a broad maximum at about 1.3 Mev. With 0.030-in. added absorber to suppress only the annihilation quanta from the N^{13} positron decay, a complete excitation curve was run. As is seen in Fig. 2, a good deal of structure was revealed. Attention is called to the scale, on which the yield at 1.76 Mev is 16 times the portion shown. The broad resonance is clearly shown. Its asymmetry is readily accounted for by the energy dependence of factors in the dispersion formula. $E_R = 1.25$ Mev and

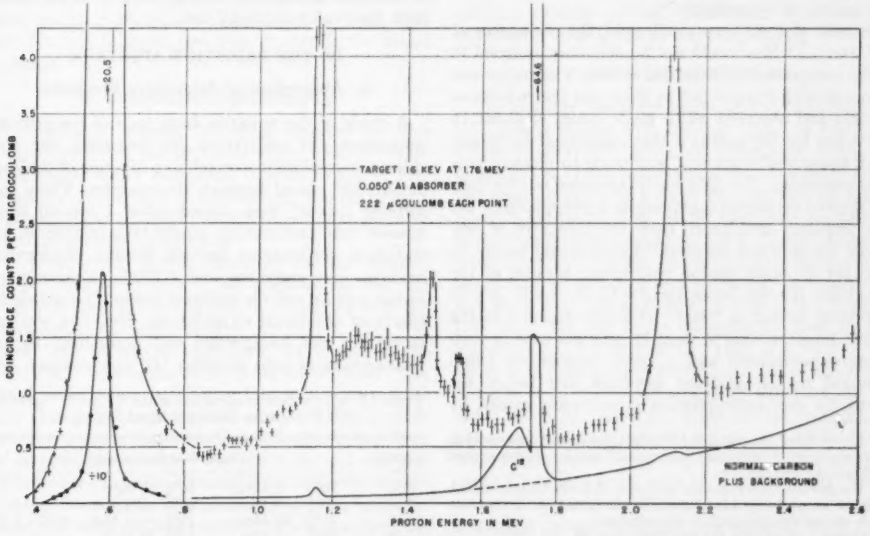


FIG. 2. Detailed thin target yield from $C^{13}(p,\gamma)N^{14}$ measured at 90° to the axis of the proton beam, showing resonances due to C^{13} at 0.55 1.16, 1.25, 1.76, and 2.10 Mev, the 1.70-Mev resonance due to C^{12}, and unidentified weak resonances at 1.47 and 1.55 Mev.

[11] Absorption in the counter walls, equivalent to 0.020 in. of aluminum, has been included in labeling the drawing "0.170 in. Al absorber."

TABLE I. Resonance energies and widths for $C^{13}(p,\gamma)N^{14}$ (in the laboratory system).

E_R (Mev)	0.55	1.16	1.25	1.76	2.10
Γ_R (kev)	32.5	6	500	2.1	45
Error (kev)	±1	±2	——	±0.2	±3

$\Gamma_R = 500$ kev are satisfactory parameters when used with the s-wave penetration factor.

The smoothed coincidence background curve is also shown in Fig. 2. Background readings were taken by bombarding a 20-kev normal carbon film on tantalum. The two targets were mounted on opposite sides of the target support, which could be rotated, so that background checks were made without delay or other changes in the system. The 1.70-Mev resonance due to C^{13} is clearly shown in the "background" curve.[12] Because the target is both thick for the C^{13} resonance and thin for the C^{12} resonance, the sharp C^{13} peak stands out clearly and indicates the width and uniformity of the target. The target enrichments were checked by the yield at 1.76 Mev, and it was found acceptable to take the manufacturer's value of 61 percent as exact. Statistically indistinguishable yields were obtained from normal films, normal lampblack, and graphite, so the stopping power of the target may safely be taken as that for pure carbon. The 0.55-Mev resonance was examined using the "Mass Two" (H_2^+) component of the beam, and the data overlapped with the Mass One data for 200 kev. Deuterium contamination of the beam was judged to be negligible.[13]

Because of their very small yield, the resonances at 1.47 and 1.55 Mev could not be definitely assigned to C^{13} by comparison with normal carbon. The region was examined with a target half as thick and the reductions in yield and observed width favor values of about 16 and 6 kev for the widths if that assignment is correct.

An unsuccessful attempt was made to identify possible impurities. The light nuclei are ruled out by their well-known excitation functions, as is phosphorus,[14] the only impurity anticipated from the chemistry of preparing the enriched material.[15] Iodine would hardly be expected to admit proton resonances, because of the forbidding barrier factor for $Z=53$. It would not be surprising to find a "trace" of iodine trapped in the target, however. One of these targets was used in scattering experiments[16] and a small unidentified bump appeared in the analyzer spectrum just before the strong Ta rise. Interpretation as scattering caused by $_{53}I^{127}$ is plausible, in which case the intensity relative to carbon scattering would imply an iodine contamination of not more than 0.5 percent.[17]

In order to obtain the true widths, the principal resonances were examined with several targets and the resonance widths calculated from the yields and observed widths. The most probable values are tabulated in Table I. In the case of the 1.76-Mev resonance, a target was prepared for which the width (from integration) was 0.5 kev. The lowest resonance was examined in somewhat greater detail, to find $Y_{max}(\xi)/Y_{max}(\infty)$, and to pursue preliminary evidence that the width was rather less than the 40 kev previously reported.[18] The resonance is also somewhat unsymmetrical, so that interpretation plays a part in the value given for Γ. The excitation function for this resonance was studied with the same target at three angles to the beam, and the target profiles were measured at 1.76 Mev for each position, so that knowledge of the target thicknesses was not dependent on accurate measurement of the target angle. The thickest target used was sufficiently thicker than the resonance width that advantage could be taken of the "infinitely" thick target yield function, Fig. 1, which was normalized in magnitude to fit the thickest "thin" target data at the point of inflection and below. $Y_{max}(\infty)$ was taken as the value at 700 kev. The internal consistency of these procedures was quite satisfactory, and led to a width $\Gamma = 32.5 \pm 1$ kev, some 20 percent lower than Fowler and Lauritsen's value[18] of 40 kev, but a resonance energy entirely in agreement with their value of 554 ± 2 kev.

IV. THE RADIATION SPECTRUM

a. Absorption of Secondary Electrons

A study of the radiation from each of the principal resonances was undertaken by measuring the coincidence counts/microcoulomb as a function of thickness of absorber placed between the counters. These "absorption curves" were carried below 1 percent transmission until background, accidental coincidences, and statistical fluctuations became serious. Background readings were measured with 0.700-in. aluminum absorber in place and the radiation present, to include the effects of accidental coincidences. The data was corrected for this background and normalized to unit transmission at zero absorber. It was analyzed with

TABLE II. Approximate composition of the $C^{13}(p,\gamma)N^{14}$ spectrum, from scintillation counter integral-bias curves.

E_R(Mev)	Character of the spectrum (γ's)
0.55	70% 8 Mev, 30% 2½ Mev
1.16	40% ~4 Mev, 60% ~2½ Mev
1.25	75% 8½ Mev, 15% ~3½ Mev, 10%<1 Mev
1.76	85% 9 Mev, 15% 2½ Mev
2.10	all 4 or 5 Mev.

[13] The enriched (61 percent) C^{13} target is of course impoverished (39 percent) in C^{12}, while the "background" is taken with a normal (98.9 percent C^{12}) fraction.

[14] The gamma-yield following (d,p) and (d,n) reactions in carbon would be at most only a few percent of the (p,γ) yield if the beam had a normal isotopic content of deuterium.

[15] Grove, Cooper, and Harris, Phys. Rev. **80**, 107 (1950).

[16] D. W. Stewart (Eastman Kodak Company), private communication. I also understand that our "normal methyl iodide," although also made by DPI, is made by a different process than the enriched material.

[17] C. W. Li and W. Whaling, Phys. Rev. **82**, 122 (1951).

[18] Ward Whaling, private communication.

[19] W. A. Fowler and C. C. Lauritsen, Phys. Rev. **76**, 314 (1949).

the aid of semi-empirical "master-curves" applicable to the geometrical arrangement used, which were prepared by interpolation[8] from the results of a study of several well-known gamma-rays. A plot of the logarithm of transmission against absorber thickness gives a family of characteristic curves of small curvature. Unfortunately, if proper account is taken of the probable errors of the data and those errors inherent in graphical solutions, this analysis cannot be carried very far. In addition, the linear dependence of detection efficiency on energy causes the high energy quanta to mask the low, and in the case of complex radiation the above method does not yield very satisfactory results. This appears to be the case for $C^{13}(p,\gamma)N^{14}$: both branching and cascades occur. However, considerable qualitative information has been obtained in this manner. Absorption curves were obtained at each of the five principal resonances, using the same target employed for the thin target yield (Fig. 2). The data at 0.55 and 1.76 Mev is very well fitted in each case by the master curve for the corresponding energy of excitation, and it is clear that the ground-state transition predominates.

In the case of the broad resonance, the tail of the absorption curve again exhibits the shape characteristic of the energy of excitation (8.7 Mev), but about 20 percent of the counts at zero absorber appear to be due to radiation of about 3 Mev, which suggests strong competition, perhaps by a branch in triple cascade.

The radiations from the 1.16- and 2.10-Mev resonances seem to be entirely of quanta less than 5 or 6 Mev, but their analysis is complicated by their weak intensity and location (see Fig. 2) astride the broad resonance.

b. Scintillation Spectrometry

Considerable interest has been aroused by this complex spectrum, but its low intensity makes conventional spectrometer techniques difficult to apply. Preliminary examination[19] of the stronger resonances by a scintillation spectrometer indicated much more sensitive detection of softer components.

E. J. Woodbury at this laboratory has used some of the C^{13}-enriched targets in an investigation[20] of the $C^{13}(p,\gamma)N^{14}$ reaction at 130 kev, in connection with the extrapolation to stellar energies. To calibrate his NaI crystal and associated equipment, we examined the $C^{13}(p,\gamma)N^{14}$ resonances and obtained integral bias curves in each case which give information to supplement that from the absorption curves. Table II gives the composition of the N^{14} radiation as determined from the integral bias curves.[20] The fractions and energies indicated are subject to systematic uncertainties similar to those involved in interpretation of the absorption curves, but with the important difference that here the soft components are not resolved less well than the hard. The results are consistent with the Geiger tube

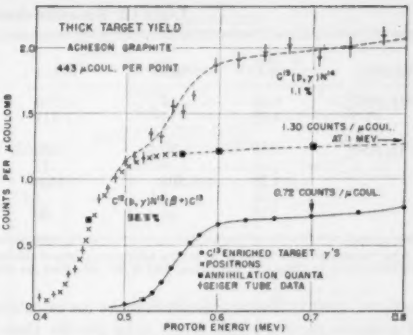


FIG. 3. Analysis of the (p,γ) yield near 0.5 Mev from a thick target of normal graphite. More accurate information on the shapes of the separate resonances has been used to disentangle the observed gamma-yield. The positron data is from reference 8 and the annihilation quanta from an entirely different measurement of the N^{13} yield, reference 2.

results, which were (linearly) weighted in favor of the hard component. It is not yet possible to make definite assignments of levels and transitions, beyond the strong suggestion that the "2½-Mev" line is the transition to the ground state from the well-known first excited state at 2.32 Mev. A very low energy contribution appears only in the broad resonance. It is to be noted that the soft components do not add up to the excitation energy unless the data represents the average energy of a complex cascade process. Additional work on this spectrum is in progress at this laboratory.‡

V. THE ABSOLUTE YIELDS

The absolute yields are based on the yield from the 0.55-Mev resonance with the thin target of Fig. 2, for which $Y_{max}=0.50Y_{max}(\infty)$, as determined during the investigation of the resonance width. The detector for this measurement was a single counter with its axis 10 cm from the target at 90° to the beam axis, surrounded by an aluminum sleeve 1.51-cm thick (to give saturation intensity of secondaries). In addition to making a direct calculation of the detection efficiency,[8] a comparison was made with the yield from a thick CaF_2 target at 1.00 Mev. For the latter we obtained 1940 counts/microcoulomb, and calculated a yield of 6.9×10^{-7} γ/p, which is in excellent agreement with the recent results of Chao *et al.*[21] The corresponding result for the C^{13} target was 8.70 counts/microcoulomb (of protons). By a similar calculation, we find a yield of 4.9×10^{-9} for a thick target taken to be 61 percent C^{13}. The yield for pure C^{13} would then be 8.0×10^{-9}, and 0.9×10^{-10} for a normal (1.12 percent C^{13}) target. It must be recognized in quoting "absolute" figures and applying small "cor-

[19] R. L. Walker and R. B. Day, Phys. Rev. **83**, 203 (A) (1951).
[20] E. J. Woodbury and W. A. Fowler, Phys. Rev. **85**, 51 (1952).

‡ *Note added in proof:* Woodbury, Day, and Tollestrup, Bull. Am. Phys. Soc. **26**, No. 8, paper J2, report more precise measurements of these γ-ray energies.

[21] Chao, Tollestrup, Fowler, and Lauritsen, Phys. Rev. **79**, 108 (1950).

TABLE III. Resonance characteristics for carbon (p,γ) reactions.[a]

Reaction	E_R (Mev)	γ_p (kev)	G (kev) s-wave	p-wave	d-wave	$Y_{max}(\infty)$ disint./proton	$\omega\Gamma_\gamma$ (ev)	σ_R (millibarns)
$C^{12}(p,\gamma)N^{13}$	0.45	35	>1750	—	—	7.6×10^{-10}	0.67	0.127
	1.70	70	117	230	1900	1.1×10^{-9}	1.39	0.035
$C^{13}(p,\gamma)N^{14}$	0.55	32.5	880	>4500	—	0.9×10^{-8}	8.6	1.44
	1.16	6	18	60	670	0.12×10^{-8}	1.3	0.56
	1.25	500	>1400	>4000	—	1.13×10^{-8}	12.8	0.062
	1.76	2.1	3.2	7.2	42	1.15×10^{-8}	14.8	12.0
	2.10	45	60	150	560	0.48×10^{-8}	6.15	1.96

[a] All energies and widths are given for the laboratory frame of reference. Yields are given for thick targets of pure isotope. Values at 1.16 and 2.10 Mev calculated from 4π times the differential yield at 90°; all others are total yields.

rections" that it is very difficult to make the calculation of absolute gamma-yields much more precise than 10 percent. The value used for the standard fluorine yield may not be much more precise.

Before enriched material was available, the thick target yield from C^{13} was estimated from the excitation function by an attempt to disentangle the "double-step." This is a very difficult experiment. With the aid of coincidence counting of the high energy radiation and positron counting of the N^{13} yield, Fowler and Lauritsen[22] were able to study the shapes of the two excitation functions separately. Both yields were found to continue to rise well "beyond" the resonance. The yields reported were 7.2×10^{-10} for C^{12} and 1.8×10^{-10} for C^{13}. The latter is twice that calculated above for the most reliable enriched target data. In view of this discrepancy, it was thought of interest to repeat the above experiment on the 3-Mev accelerator. By the use of the copious H_2^+ components of the beam, equivalent to 4 microamperes of protons, it was possible to obtain data with a 3 percent standard deviation in bombardments lasting only 2 minutes, permitting the contribution of annihilation quanta to be included in background measurements alternating with bombardment. The results are shown in Fig. 3. In view of the serious statistical scatter of the earlier data (not shown), it is in agreement with the present excitation function. The double excitation curve was analyzed with the use of Fowler and Lauritsen's positron curve, and the gamma-curve (Fig. 1) obtained with enriched material. Detailed interpretation of the curve with the advantage of improved hindsight about the separate shapes gives considerably different results. CaF_2 was again used as an independent check on efficiency calculations, and the yields calculated are 7.0×10^{-10} for C^{12} (at 1.00 Mev) and 1.1×10^{-10} for C^{13} (at 700 kev), in normal graphite.

The data obtained in calibrating Woodbury's counter on CaF_2 and carbon targets permits calculation[23] of the yield from C^{13} in normal graphite as 0.9×10^{-10} γ/p for the hard component, and a total yield of 1.0×10^{-10} N^{14}/p if the soft components are regarded as three equal quanta in cascade.

An additional check on the ratio of yields due to C^{12} and C^{13} near 0.5 Mev is available from the ratio near 1.7 Mev, but at present it offers no increase in precision. We conclude that the yield at 700 kev for a thick target of pure C^{13} is 0.9×10^{-8} disintegration/proton. This value supersedes earlier measurements made at this laboratory and has a probable error about 10 percent in excess of that of the fluorine yield used as a reference.

VI. WIDTHS AND CROSS SECTIONS

From the measured proton widths Γ_p and the value of $Y_{max}(\infty)$ for the 0.55-Mev resonance, the values of $Y_{max}(\infty)$ have been calculated for each of the other resonances using the data of Fig. 2 and the absorption curves. The width for gamma-emission cannot be obtained directly, but the product $\omega\Gamma_\gamma$ can readily be calculated, where ω is a statistical factor. For a target of C^{13}, $\omega=(2J+1)/4$, where $\hbar J$ is the total angular momentum of the compound state. The cross sections at resonance are on the order of millibarns, and are numerically related to the widths and resonance energies (in the laboratory frame of reference) by the expression

$$\sigma_R=0.301(1\,\text{Mev}/E_R)(\omega\Gamma_\gamma/1\,\text{ev})(10\,\text{kev}/\Gamma_p)\ \text{millibarns}.$$

Table III gives the resonance characteristics for each of the resonances studied and the two resonances in $C^{12}(p,\gamma)N^{13}$. Also tabulated are the "nuclear widths without barrier." G_l calculated from the curves of Christy and Latter[24] for s-, p-, and d-wave protons, which give a measure of the dependence of widths on intrinsically nuclear factors. R. G. Thomas has shown[25] that estimation of the nuclear width in this manner may be in error for broad resonances, so that the larger values (over 1 Mev) in the table should be somewhat larger.

VII. DISCUSSION

Without better information on the spectrum very few unequivocal conclusions can be drawn about the structure of the nuclear states involved. It appears that the 0.45-Mev resonance with C^{12} and the 0.55- and 1.25-Mev resonances with C^{13} are formed by capture of s-wave protons, since a nuclear width of a few Mev corresponds

[22] This data is presented in Fig. 1 of reference 8.

[23] E. J. Woodbury (private communication). See reference 20.

[24] R. F. Christy and R. Latter, Revs. Modern Phys. **20**, 185 (1948).

[25] R. G. Thomas, Phys. Rev. **80**, 136 (A) (1950), and Ph.D. thesis, California Institute of Technology (1951). A paper on this subject is forthcoming.

to decay in a time comparable to the time of formation of the compound state. More refined calculations of "reduced widths" for the broad resonances, based on a modified dispersion theory[25] indicate that this is indeed so. Moreover, the two resonances in N^{14} have nearly equal reduced widths, and the large level spacing predicted for such levels requires that these two have different "spin,"[26] i.e., $J=0$, or $J=1$. The ground state of N^{14} has $J=1$, and presumably even parity.[27] For both of these resonances, the observed radiation widths correspond to reasonable electric dipole matrix elements for either $J=0$, or $J=1$.

Assignment at $J=0$ to the $E_p=0.55$-Mev level would forbid the 5.8-Mev direct transition to the first excited state of 2.3 Mev, which is taken to have $J=0$ and even parity, as it is presumably the analog of the C^{14} and O^{14} ground states (see below). The corresponding transition from the broad level ($J=1$) should then be allowed, but the present evidence seems to indicate that the branch competing with the ground-state transition consists of quanta of lower energy.

Another competing mechanism which may be suggested is the possibility of magnetic radiation between the two broad resonances, followed by ejection of a proton from the lower state, since the proton width is several thousand times the radiation width. Inelastically scattered protons should be observable.

For the 1.16-, 1.76-, and 2.10-Mev resonances, the observed widths correspond to reasonable matrix elements for electric dipole radiation of the energy observed, but very little more can be said. The very small width and anisotropic radiation observed by Day exclude s-wave formation of the resonance at 1.76 Mev, and an assignment of $J=2$, odd parity is consistent with all results, but the assignment is not unique.

Figure 4 is an energy-level diagram for N^{14}, showing the levels now known and the regions explored by this and other work. The reader is referred to reference 1, page 335 for a complete diagram, including the levels above 11 Mev. It is clear that our cascades could take place in a number of ways. The region of excitation between 9.5 and 10.5 Mev remains unexplored and requires energies up to 3.2 Mev for the (p,γ) reaction. Also the region between 6.7 and 7.7 Mev is inadequately explored. With the advent of thin C^{13} targets, the $C^{13}(d,n)N^{14}$ reaction may be used, and the yield of slow (threshold) neutrons investigated. An investigation of proton scattering by C^{13} is being undertaken at this laboratory and is expected to give information on the values of J for the levels found in the (p,γ) reaction.

Also shown in Fig. 4 are the ground states of O^{14} and C^{14}. These occur at the energies shown in parentheses, but when the Coulomb energy and n^1-H^1 differences are removed, they correspond very well to the first excited state of N^{14}. A level at 6.11 Mev in C^{14} has been found in the $C^{13}(d,p)C^{14}$ reaction.[28] This state must have

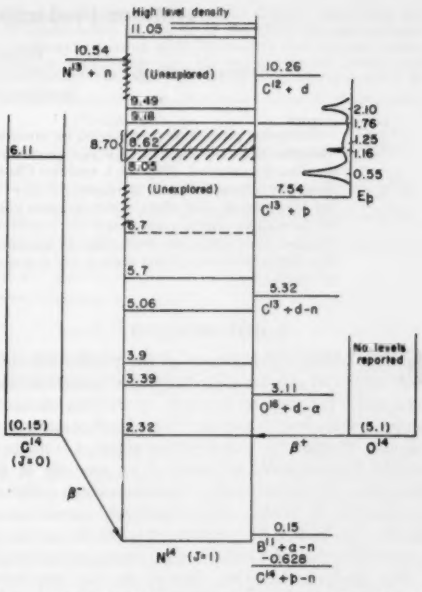


FIG. 4. Energy-level diagram for N^{14} and related nuclei. The Q-values given are based on the recent results of Li, Whaling, Fowler, and Lauritsen, Phys. Rev. **83**, 512 (1951).

$J\neq 0$ to radiate as observed to the ground state ($J=0$), and has presumably $J=1$ and odd parity. This state may be the analog of the level in N^{14} at 8.05 Mev.[29] If this is correct, the 8.05-Mev level in N^{14} must (also) have $J=1$, odd parity, and the 8.70-Mev level have $J=0$, odd parity, but a more complex explanation is required for the apparent failure of the 8.05-Mev level to radiate to the 2.3-Mev level.

VIII. CONCLUSION

Five states in N^{14} have been investigated, and shown to emit a complex radiation spectrum, suggesting and challenging further experimental and theoretical work. The next important step should be an intensive attack on the energies, intensities, and branching ratios of the spectrum. It is hoped that the new techniques of scintillation spectroscopy may prove a useful tool in that investigation. The properties of more levels in regions of excitation not yet investigated may supply additional information.

The author is indebted to R. B. Day and J. E. Perry, Jr. for their collaboration in the early phases of this work, and to E. J. Woodbury for his measurements of the yield and spectrum. It is a pleasure to acknowledge the guidance and interest of the entire staff of the Kellogg Radiation Laboratory through this investigation.

[25] R. G. Thomas (private communication).
[26] S. T. Butler, Phys. Rev. **80**, 1095 (1950).
[27] R. G. Thomas and T. Lauritsen, Phys. Rev. **78**, 88 (1950).
[28] R. G. Thomas (private communication).

π-Meson Production by Protons on Nuclei*

ERNEST M. HENLEY
Physics Department, Radiation Laboratory, University of California, Berkeley, California
(Received September 24, 1951)

The meson production cross section for protons incident on nuclei is formulated in terms of two-particle transition rates and the struck nucleon momentum distribution in the nucleus. Three different momentum densities are assumed. They are a modified Chew-Goldberger, a Gaussian, and a Fermi distribution. With these, it is attempted to fit the experimental π^+ and π^- meson spectra obtained by bombarding C^{12} with 345-Mev protons. The effect of the exclusion principle is estimated on the basis of a single particle model for the nucleus. Meson reabsorption and scattering, after production, are also taken into account, but only roughly. The calculations show that the gaussian distribution approximates nuclear conditions best, and that the fundamental proton-neutron and proton-proton production transitions matrices are most probably not equal.

I. INTRODUCTION

EXPERIMENTAL studies of the production spectrum and angular distribution for π-mesons have been made for protons incident on various elements, notably hydrogen,[1-3] carbon,[4,5] and lead[6] at energies of 345 and 381 Mev. In this paper an attempt is made to explain these spectra in terms of an analysis of the production of mesons in free nucleon-nucleon collisions as described by Watson and Brueckner,[7] (hereafter referred to as W.B.). The treatment is similar to that of Lax and Feshbach for photomeson production in nuclei.[8]

The analysis need not depend on any particular meson theory, but assumes that the matrix element for meson production in proton-proton ($P-P$), and proton-neutron ($P-N$) collisions are known in detail. Unfortunately, little is presently known about the latter, nor is the excitation curve known for either types of meson production. On the other hand, a great deal of evidence (see reference 7 for example) seems to indicate that both the charged and neutral mesons are pseudoscalar. The spin dependence of the matrix for the production from either proton-proton or proton-neutron collisions is based on this assumption, but the calculation can easily be extended to apply to any other type of meson.

The evaluation of the transition matrix of the problem is carried out by performing closure over all but the interacting nucleons. Except near threshold, the production spectrum at a given angle of meson emission is then shown to be proportional to the momentum distribution of the struck particle in the nucleus, folded into the free particle production rate. The consequence of the exclusion principle is to reduce the momentum space available, and is calculated on the basis of a single particle model for the nucleus.

The absorption and scattering of the meson by the nucleus, after production, are treated as separate processes.

At sufficiently high energies above threshold, the cross section is expected to be equal to the number of protons and the number of neutrons in the nucleus multiplied by their respective free-particle production cross sections. As the energy of the incident proton is decreased, not all the nucleons can contribute to the process (i.e., the whole momentum distribution is not available), and the "production efficiency" may be said to decrease, approaching zero at threshold. This efficiency is a consequence of the dynamics of the problem and is independent of the exclusion effects, which will tend to decrease the cross section still further near threshold. At a given proton energy, the efficiency is a function of the angle of emission of the meson. It is largest for mesons emitted forward. For mesons emitted at 90° to an incident proton beam of 341 Mev, it is only somewhat over 0.5.

The theoretical results derived in the first part are applied to carbon bombarded by 341-Mev protons. It is shown that the spectrum obtained for meson production at 90° to the incident beam depends mainly on the nuclear momentum density. At 0°, however, for meson energies up to 80 Mev, the spectrum depends largely on the free particle meson production matrices. For the sake of simplicity these are assumed to differ only by a numerical constant for $P-N$ and $P-P$ collisions, even though this hypothesis will be shown to be inadequate. It is then attempted to deduce some information both about the matrix elements and about the nuclear momentum distribution involved in the cross section.

II. FORMULATION

A. Free Particle Meson Production

Before considering the actual problem of meson production in nuclei, it is worthwhile to review some of the features of creation in free particle interactions. The cross section for meson formation in a nucleon-nucleon collision can be expressed in terms of an

* This work was performed under the auspices of the AEC.
[1] Cartwright, Richman, Whitehead, and Wilcox, Phys. Rev. **78**, 823 (1950).
[2] Peterson, Iloff, and Sherman, Phys. Rev. **81**, 647(A) (1951).
[3] M. N. Whitehead and C. Richman, Phys. Rev. **83**, 855 (1951).
[4] C. Richman and H. A. Wilcox, Phys. Rev. **78**, 496 (1950)—345 Mev.
[5] Block, Passman, and Havens, Phys. Rev. **83**, 167 (1951)—381 Mev.
[6] M. Weissbluth, Phys. Rev. **78**, 86(A) (1950).
[7] K. M. Watson and K. A. Brueckner, Phys. Rev. **83**, 1 (1951).
[8] M. Lax and H. Feshbach, Phys. Rev. **81**, 189 (1951).

R-matrix.[7,9] W.B. show that, due to the short range of interaction required for meson production, a zero range approximation can be introduced for the nucleons in the final state. If the meson is assumed to come off as a plane wave, and momentum conservation is factored out (in units of $\hbar=c=1$)

$$d\sigma=[(2\pi)^4/v_R]|(2\pi)^{\frac{1}{2}}\chi_{p'}(0)|^2|(\mathbf{q}',\mathbf{p}'|r|\mathbf{p})|^2 \times\delta(\mathbf{q}+\mathbf{n}_1'+\mathbf{n}_2'-\mathbf{n}_1-\mathbf{n}_2)d\mathbf{J}, \quad (1)$$

where $\mathbf{q}$ is the meson momentum, $\mathbf{p}'$ and $\mathbf{p}$ the final and initial relative momenta, respectively, and v_R the relative velocity of the interacting nucleons. If μ, and M are the meson and nucleon masses, $\mathbf{n}_1$ and $\mathbf{n}_2$ the individual particle momenta,

$$\mathbf{q}'=\mathbf{q}-(\mu/2M)(\mathbf{n}_1+\mathbf{n}_2),$$
$$\mathbf{p}'=\tfrac{1}{2}(\mathbf{n}_1'-\mathbf{n}_2'),$$
$$\mathbf{p}=\tfrac{1}{2}(\mathbf{n}_1-\mathbf{n}_2).$$

The $|(2\pi)^{\frac{1}{2}}\chi_{p'}(0)|^2$ factor in Eq. (1) ($\mathbf{p}'$ refers to the eigenvector and 0 is used to indicate the zero-range approximation) comes from the final state nucleon interaction, on the hypothesis that only their relative motion S-state need be considered.[10] If q_0 is the meson energy, T_F and T_I the final and initial nucleon kinetic energies, respectively, then the phase space volume available,

$$d\mathbf{J}=d\mathbf{q}d\mathbf{n}_1'd\mathbf{n}_2'\delta(q_0+T_F-T_I).$$

A sum over the final spin states and an average over the initial ones is implied in Eq. (1).

A partial wave analysis of the final nucleon states has been used. If a similar consideration is applied to the meson angular momentum, and only P- and S-states are included, the matrix $|r|^2$ in the cross-section formula can be written as:

$$|r^P|^2=a(q^2/\mu^2)\cos^2\theta+b(q^2/\mu^2)+e, \quad (2a)$$

for meson production in $P-P$ collisions, and as

$$|r^N|^2=a'(q^2/\mu^2)\cos^2\theta+b'(q^2/\mu^2) +c'(q^2/\mu^2)\cos\theta+d'(q/\mu)\cos\theta+e', \quad (2b)$$

for mesons produced in $P-N$ collisions. The coefficients, a, b, e, etc., are numerical factors which depend on the initial momentum, $\mathbf{p}$, and on the spin and isotopic spin states of the reaction. θ is the meson angle in the center-of-mass system of the interacting particles. No $\cos\theta$ terms appear in Eq. (2a) because of the exclusion principle.

W.B. investigated the consequences of charge independence for nuclear forces, taken together with angular momentum and parity conservation, as well as the Pauli exclusion principle, for the process considered. For mesons created by incident protons their results are summarized in Table I. The case of π^0 production will not be studied in any further detail here, though the treatment can easily be extended to include it as well.

TABLE I. Transitions permitted by angular momentum and parity conservation, as well as by the Pauli exclusion principle for pseudoscalar mesons, when the relative final nucleon angular momentum is an S-state. M_1, M_2, and M_3 are the isotopic spin representation of the R-matrices. The notation used is that of W.B.

	Meson in S-state		Meson in D-state	
$P+P$, π^+	${}^3P_0\to{}^1S_0$ $M_1=(1^+s\|r^P\|s^-)$ ${}^1P_1\to{}^3S_1$ $M_3=(1^+s\|r^P\|s^-)$		${}^1S_0\}\to{}^1S_1$ $M_3=(1^+s\|r^P\|s^-)$ ${}^1D_2\}$	
$P+N$, π^+	${}^3P_0\to{}^1S_0$ $M_1=(1^+t^+\|r^N\|s^0)$		${}^3S_1\}\to{}^1S_0$ $M_2=(1^+t^+\|r^N\|s)$ ${}^3D_1\}$	
$P+N$, π^-	${}^3P_0\to{}^1S_0$ $M_1=-(1^-t^-\|r^N\|s^0)$		${}^3S_1\}\to{}^1S_0$ $M_2=(1^-t^-\|r^N\|s)$ ${}^3D_1\}$	
$P+P$, π^0[a]	${}^3P_0\to{}^1S_0$ $M_1=-(1^0t^-\|r^N\|s^-)$		forbidden	$M_1=-(1^0t^-\|r^P\|s^-)$
$P+N$, π^0	${}^3P_0\to{}^1S_0$ 0		${}^3S_1\}\to{}^1S_1$ 0 ${}^3P_{1,2}\}$	
	${}^3P_1\to{}^1S_1$ $M_2=(1^0s\|r^N\|s^0)$		${}^3S_1\}\to{}^1S_0$ $M_2=(1^0s^0\|r^N\|s)$ ${}^3D_1\}$	
	${}^1P_1\to{}^3S_1$ 0		${}^1S_0\}\to{}^3S_1$ $M_3=(1^0s\|r^N\|s^0)$ ${}^1D_2\}$	

[a] Bjorkland, Crandall, Moyer, and York, Phys. Rev. **77**, 213 (1950) show experimental evidence of the suppression of this reaction.

W.B. also obtained the general spin dependence of the R-matrices on the assumption of pseudoscalar mesons. If the index "1" refers to one of the interacting particles and "2" to the other, this dependence is given by a combination of

$$A(1)=\mathbf{T}(\mathbf{p},\mathbf{q})\cdot\boldsymbol{\sigma}^{(1)}+\mathbf{U}(\mathbf{p},\mathbf{q})\cdot\boldsymbol{\sigma}^{(1)}\mathbf{V}(\mathbf{p},\mathbf{q})\cdot\boldsymbol{\sigma}^{(2)} +\mathbf{W}(\mathbf{p},\mathbf{q})\cdot\boldsymbol{\sigma}^{(1)}\times\boldsymbol{\sigma}^{(2)}, \quad (3)$$

and similar quantities $A(2)$ obtained by interchanging "1" and "2" and replacing $\mathbf{p}$ by $-\mathbf{p}$. The R-matrix must be symmetric with respect to nucleon exchange. Thus, for a π^+ meson P-state, if T is an even function of $\mathbf{p}$, we would take $\mathbf{T}\cdot(\boldsymbol{\sigma}^{(1)}-\boldsymbol{\sigma}^{(2)})$, since the isotopic spin dependence (see Table I) of R is odd.

In this manner, then, except for numerical factors, an almost complete specification of the free particle production matrix is possible.

B. Production in Nuclei

For meson production in a fairly heavy nucleus, at energies of the order of 350 Mev, the interaction can be assumed to take place between the incoming proton and a single nucleon. The wavelength of the incoming proton, as seen by the nucleus (if $A\gg1$, the center-of-mass system and the laboratory frame are approximately equivalent) is $\lambda=2.3\times10^{-14}$ cm. Since this is quite small compared to internuclear distances ($\sim2.8\times10^{-13}$ cm) the above approximation is justified. On the basis of an impulse approximation[11] then, the probability amplitude for meson production in a nucleus, A, is proportional to

$$\left(G_F,\sum_{j=1}^{A}R_jG_I\right),$$

[9] C. Møller, Kgl. Danske Videnskab. Selskab, Mat.-fys. Medd. **23**, 1 (1945).

[10] See Appendix, K. Bruckner, Phys. Rev. **82**, 598 (1951).

[11] G. F. Chew, Phys. Rev. **80**, 196 (1950).

where R_j is the appropriate two-particle matrix of the form

$$(\mathbf{q}, \mathbf{n}_1', \mathbf{n}_2'|R|\mathbf{n}_1, \mathbf{n}_2).$$

Momentum conservation can be factored out, giving[12]

$$(\mathbf{q}', \mathbf{p}'|r|\mathbf{p})\delta(\mathbf{q}+\mathbf{n}_1'+\mathbf{n}_2'-\mathbf{n}_1-\mathbf{n}_2).$$

G_F and G_I are the final and initial antisymmetrized nuclear wave functions in momentum space. Thus, if the momentum, $\mathbf{n}''$, spin and isotopic spin, ξ, of a particle, i, is indicated by η_i'',

$$G_I=\frac{1}{(A+1)!}\sum_{i=1}^{A+1}(-1)^P P_{i,A+1} \times\psi_I(\eta_1''\cdots\eta_i''\cdots\eta_A'')\epsilon(\eta_{A+1}''), \quad (4)$$

where ψ_I is antisymmetric and $(-1)^P P_{i,A+1}$ is the particle permutation operator which insures that G_I is also antisymmetric. ϵ is the wave function of the incoming particle. This is a plane wave, and may be written as

$$\epsilon(\eta_{A+1}'')=\delta_{\mathbf{P}}(n_{A+1}'')\epsilon'(\xi_{A+1}),$$

where ϵ' represents the spin, isotopic spin wave function.

The cross section for meson creation is then

$$d\sigma_A=\frac{(2\pi)^4}{v_0}\sum_F\left|\left(G_F,\sum_{j=1}^{A}R_jG_I\right)\right|^2\delta(E_F-E_I)d\mathbf{q}d\mathbf{P}', \quad (5)$$

where v_0 is the relative velocity of the incident proton and the struck nucleus, $\mathbf{P}'$ is the total momentum of the interacting nucleons in the final state, and $\sum_F$ indicates a sum over all possible final nuclear states, consistent with over-all energy conservation.

Since the meson-producing interacting is a strong, short-range one, it is expected that the excitation energy of the final nucleus, excluding the two colliding particles will be small. If the effect of this slight excitation on the energy conservation is neglected, we can sum over all final states of $(A-1)$ particles, and thus obtain a partial closure approximation to the cross section. The exact energy conservation condition is then replaced by:

$$\delta(T_F+q_0+\bar{B}_F-T_I-B_I), \quad (6)$$

where T_F is the final kinetic energy of the interacting nucleons, and T_1 the initial kinetic energy of the incident proton, since the initial nucleus is at rest. B_I and $\bar{B}_F$ are the initial and average final nuclear binding energies, respectively.

Due to the large momentum transfer involved in the interaction, we can take the final state wave function to be separable. Then, if ψ_F and χ are antisymmetric wave functions

$$G_F(\eta_1\cdots\eta_{A+1})=\frac{1}{A}\sum_{r\neq s}\sum_{s\neq r}(-1)^P P_{A,r}\,{}_{A+1,s} \times\psi_F(\eta_1\cdots\eta_r\cdots\eta_s\cdots\eta_{A-1})\chi(\eta_A,\eta_{A+1}). \quad (7)$$

[12] In order not to obtain the square of the δ-function, the artifice employed is that of using Q instead of q in the second δ-function, and integrating over this momentum.

With this separation, the evaluation of the matrix involved in the cross section Eq. (5) is carried out as follows:

$$M=\sum_F|(G_F,\sum_j R_jG_I)|^2\equiv\sum_F|(G_F,\phi_I)|^2. \quad (8)$$

This defines $\phi_I\equiv\sum_j R_jG_I$, which must also be antisymmetric. If closure is performed over all final states of ψ_F, and a sum over the final spins of the struck nucleons, M is given by

$$M=\frac{1}{A^2}\sum_{r\neq s}\sum_{s\neq r}{}'\langle\eta_1\cdots\eta_r'\cdots\eta_s'\cdots\eta_{A+1} \times|\chi^*(\mathbf{n}_r,\mathbf{n}_s)\chi(\mathbf{n}_r',\mathbf{n}_s')\delta(\xi_r'-\xi_r)\delta(\xi_s'-\xi_s)| \times\eta_1\cdots\eta_r'\cdots\eta_s'\cdots\eta_{A+1}\rangle', \quad (9)$$

where $|\eta_1\cdots\eta_{A+1}\rangle'$ represents ϕ_I, and an integration over all variables is implied. All terms in the sum are similar. The matrix may thus be replaced by a typical element

$$M={}'\langle\eta_1'\eta_2'\eta_3\cdots\eta_{A+1}|\chi^*(\mathbf{n}_1,\mathbf{n}_2)\chi(\mathbf{n}_1',\mathbf{n}_2') \times\delta(\xi_1'-\xi_1)\delta(\xi_2'-\xi_2)|\eta_1\eta_2\eta_3\cdots\eta_{A+1}\rangle'.$$

If ψ_I is represented by $|\eta_1\cdots\eta_A\rangle$, and if the free particle R-matrices are given an added index to indicate both interacting particles, the meson production matrix becomes

$$M=\frac{1}{A+1}\sum_{i\neq i}\sum_i\sum_{l\neq k}\sum_k\langle\eta_1'\eta_2'\eta_3\cdots\eta_i'''\cdots\eta_A| \times\epsilon^*(\eta_k''')R_{kl}\chi^*(\mathbf{n}_1,\mathbf{n}_2)\chi(\mathbf{n}_1',\mathbf{n}_2')\delta(\xi_1'-\xi_1) \times\delta(\xi_2'-\xi_2)R_{ij}\epsilon(\eta_i'')|\eta_1\eta_2\cdots\eta_j''\cdots\eta_A\rangle.$$

C. Diagonal Elements

The matrix M has both diagonal and off-diagonal elements. The diagonal terms represent the main contribution to the cross section, except near threshold, since the off-diagonal elements would not occur at all if correlation effects did not exist for the location of nucleons and if wave interference is neglected. The diagonal terms, M^0 are considered first. If the center-of-mass motion of $\chi(\mathbf{n}_i,\mathbf{n}_j)$ is factored out

$$M^0=\frac{1}{A+1}\sum_{j\neq i}\sum_i\langle\psi_I|\epsilon'^*(\xi_i')R_{ij}\delta_{\mathbf{P}'}(\mathbf{n}_i+\mathbf{n}_j) \times\chi^*(\mathbf{n}_i-\mathbf{n}_j)\delta_{\mathbf{P}}(\mathbf{n}_i'+\mathbf{n}_j')\chi(\mathbf{n}_i'-\mathbf{n}_j')R_{ij}\epsilon'(\xi_i)|\psi_I\rangle. \quad (10)$$

Though the colliding nucleons may have final momenta considerably larger than in the free particle case, it can be argued, that due to their intimate interaction, the zero range hypothesis of W.B. can still be used. The final state of these nucleons is again approximated by considering relative angular momentum S-waves only. $\chi(\mathbf{n}_i-\mathbf{n}_j)$ is thus replaced by $\chi_{P'}(0)$ for S-states, and is

taken as zero otherwise. With this simplification, and if momentum conservation is factored out,

$$M^0=\frac{1}{A+1}\sum_{j\neq i}\sum_i\langle\psi_I|\epsilon'^*(\xi_i')|r_{ij}|^2|(2\pi)^{\frac{1}{2}}\chi_{\mathbf{p}'}(0)|^2 \times\delta(\mathbf{q}+\mathbf{P}'-\mathbf{P}-\mathbf{k})\epsilon'(\xi_i)|\psi_I\rangle, \quad (11)$$

where $\mathbf{k}$ is the struck nucleon momentum eigenvector.

The spin and isotopic spin sums in Eq. (11) can be performed by making use of the representation of the R-matrices shown in Sec. A. Thus, by means of Table I, the isotopic spin dependence of M^0 is completely specified. For mesons emitted in a P-state, (omitting the δ-function and the factor $|(2\pi)^{\frac{1}{2}}\chi_{\mathbf{p}'}(0)|^2$ for the moment),

$$M^0(\pi^+)=\langle\psi_I|\sum_{j\neq i}\sum_i\{(1^+s|r_{ij}{}^P|t^-)^2 +(1^+t^+|r_{ij}|s)^2\}|\psi_I\rangle, \quad (12)$$

$$M^0(\pi^-)=\langle\psi_I|\sum_{j\neq i}\sum_i(1^-t^-|r_{ij}{}^N|s)^2|\psi_I\rangle. \quad (13)$$

Similar expressions are obtained if the meson is emitted in an S-state. The actual spin dependence, as shown in Eq. (3) is not required. It can be shown that, for the transitions involved (see Table I) we need only consider

$$r_{ij}{}^P=\mathbf{K}\cdot\boldsymbol{\sigma}_j \quad \text{and} \quad r_{ij}{}^N=\mathbf{L}\cdot\boldsymbol{\sigma}_j,$$

operating on the initial nucleus wave function, where $\mathbf{K}$ and $\mathbf{L}^*$ are functions of the meson and colliding nucleon momenta. All other forms cancel for pseudoscalar mesons.

Averaging over all nuclear spins, and performing the isotopic spin sum, we obtain for both meson P- and S-states

$$M^0(\pi^+)=\left\langle\psi_I\left|ZK^2+\frac{A-Z}{2}L^2\right|\psi_I\right\rangle, \quad (14)$$

$$M^0(\pi^-)=\left\langle\psi_I\left|\frac{A-Z}{2}L^2\right|\psi_I\right\rangle, \quad (15)$$

where Z is the number of protons, and A the total number of nucleons in the nucleus.

These matrix elements can be expressed in terms of the normalized momentum distribution of the struck nucleon, $\rho(\mathbf{k})$ by integrating over all nuclear coordinates on which the R-matrix does not depend.

The contribution of the diagonal elements to the cross section is thus given by (if the omitted factors are again taken into account):

$$d\sigma_A{}^0(\pi^+)=\frac{(2\pi)^4}{v_0}\int\left(ZK^2+\frac{A-Z}{2}L^2\right) \times|(2\pi)^{\frac{1}{2}}\chi_{\mathbf{p}'}(0)|^2\rho(\mathbf{k})d\mathbf{k}\delta(\mathbf{q}+\mathbf{P}'-\mathbf{k}-\mathbf{P}) \times\delta(E_F-E_I)d\mathbf{P}'d\mathbf{p}'d\mathbf{q}, \quad (16)$$

since $d\mathbf{q}=q_0qdTd\Omega_q$, where T is the meson kinetic energy and $d\Omega_q$ is an element of solid angle about the direction of emission of the meson, we have, in terms of the free particle cross sections

$$\frac{d\sigma_A{}^0(\pi^+)}{dTd\Omega_q}=\frac{1}{v_0}\int\left[Z\frac{d\sigma_{P-P}}{dTd\Omega_q}v_R +(A-Z)\frac{d\sigma_{P-N}}{dTd\Omega_q}v_R\right]\rho(\mathbf{k})d\mathbf{k}, \quad (17a)$$

$$\frac{d\sigma_A{}^0(\pi^-)}{dTd\Omega_q}=\frac{1}{v_0}\int(A-Z)\frac{d\sigma_{P-N}}{dTd\Omega_q}v_R\rho(\mathbf{k})d\mathbf{k}. \quad (17b)$$

The correct energy conservation condition, as given by Eq. (6), is implied in these equations, and it is assumed that the excitation function of the free particle cross section is known.

On the basis of the diagonal terms alone, if $K^2=L^2$, the ratio of

$$d\sigma_A{}^0(\pi^+)/d\sigma_A{}^0(\pi^-)=(A+Z)/(A-Z).$$

For a nucleus such as carbon, where $Z=A/2$, this ratio is 3:1.

D. Correlation Effects

It remains to consider the off-diagonal elements, $\bar{M}$, of the matrix M. These occur because of wave interference effects and due to the antisymmetrization of the wave functions. Since the production process involves large momentum transfers, the former type is expected to be small. It is neglected in the actual calculations, where a single-particle model is used to evaluate the correlation terms. For the present, however, these restrictions will not be imposed.

The magnitude of the matrix $\bar{M}$ decreases as the energy of the incident proton increases above threshold. Since the latter is at approximately 165 Mev,[13] $\bar{M}$ is not expected to alter the cross section appreciably at 341 Mev.

If an average is performed over the spins of the incoming particle, the general off-diagonal matrix element, referring to Eq. (9), can be written as

$$\bar{M}=\frac{1}{A+1}\sum_{l\neq j\neq i}\sum_{j\neq i}\sum_i\langle\eta_1'\eta_2'\eta_3\cdots\eta_l'''\cdots\eta_A|\epsilon'^*(\xi_i) \times R_{il}{}^{\frac{1}{2}}\chi^*(\mathbf{n}_1,\mathbf{n}_2)\chi(\mathbf{n}_1',\mathbf{n}_2')\delta(\xi_1'-\xi_1) \times\delta(\xi_2'-\xi_2)R_{ij}\epsilon'(\xi_i)|\eta_1\cdots\eta_j''\cdots\eta_A\rangle. \quad (18)$$

$\bar{M}$ contains several types of terms, corresponding physically to an exclusion of either or both of the interacting particles in the final state. With reference to Eq. (18) the three distinct off-diagonal matrices which occur may be represented by:

(1) M'; this occurs if $i=1$; $l\neq j\neq 1$ or 2;

(2) M''; this occurs if $j=1$; $l\neq i\neq 1$ or 2;

(3) M'''; this occurs if $l\neq j\neq i\neq 1$ or 2.

[13] W. H. Barkas, Phys. Rev. 75, 1109 (1949). He obtains 155 Mev for π^+ production and 178 Mev for π^- production.

M' and M'' involve an overlap integral over a single particle variable, whereas M''' involves one over two variables. If these overlaps are large, the contribution of the off-diagonal matrices will be important, and in a direction to cancel the diagonal ones.

The spin and isotopic spin sums in M' and M'' can be evaluated as by Lax and Feshbach[2] by methods due to Wigner and Feenberg.[14] M', for example, involves pair correlation functions $\rho(\mathbf{n}_j, \mathbf{n}_l)$. To a first approximation only momentum space symmetric and antisymmetric correlations, $\rho(|\mathbf{n}_j-\mathbf{n}_l|)$ are differentiated. The actual dependence of M' on $\mathbf{n}_l$ and $\mathbf{n}_j$ can only be obtained if χ is known, and if the overlap is performed. If $\chi(\mathbf{n}_i, \mathbf{n}_2)$ is replaced by individual plane waves:

$$\chi(\mathbf{n}_i, \mathbf{n}_2)=\delta_{p2}(\mathbf{n}_2)\delta_{p_i}(\mathbf{n}_i),$$

we obtain for M'

$$M'=\frac{1}{A+1}\sum_{l\neq j\neq i}\sum_{j\neq i}\sum_{i}\langle\psi_I|\epsilon'(\xi_i)r_{il}{}^{\dagger}r_{ij}\delta(\mathbf{n}_j''-\mathbf{n}_l''')\times\delta(\mathbf{P}-\mathbf{q}-\mathbf{p}_i)\epsilon(\xi_i)|\psi_I\rangle. \quad (19)$$

The exclusion effect is now implicit in the wave function ψ_I, as will be shown in more detail, later. A similar result is obtained for M'', except that the final momentum involved in the δ-function is now that of the other nucleon in the final state, $\mathbf{p}_2$.

If V_{jl} is defined as the momentum space orientation mean value of $\delta(\mathbf{n}_j''-\mathbf{n}_l'')\delta(\mathbf{P}-\mathbf{q}-\mathbf{p}_i)$ and an average over nuclear spins is performed, we obtain for mesons emitted in P-states:

$$M'(\pi^-)=\frac{1}{A+1}\sum_{l\neq j\neq i}\sum_{j\neq i}\sum_{i}\langle\psi_I|\epsilon'^{*}(\xi_i)(L^2\boldsymbol{\sigma}_j\cdot\boldsymbol{\sigma}_l)\times(s_{il}|0|1^-t^-)(1^-t^-|0|s_{ij})\epsilon(\xi_i)|\psi_I\rangle, \quad (20)$$

where 0 is the isotopic spin representation of the R-matrices, and the notation is that of Table I. A similar, but more complicated result is obtained for $M'(\pi^+)$, since contributions occur from both $P-P$ and $P-N$ collisions, though there is no interference between these two since their 0-operators are orthogonal (M_2, M_3) in isotopic spin space. For mesons emitted in S-states, such interference is possible as can be seen by referring to Table I.

The state $|\psi_I\rangle$ is now broken up into its symmetric and antisymmetric parts $|\psi_I\rangle=|\psi_I{}^{(s)}+|\psi_I{}^{(a)}\rangle$ and the spin and isotopic spin sums are carried out. If

$$V_s\equiv\langle\psi_I{}^{(s)}|V_{jl}|\psi_I{}^{(s)}\rangle \quad \text{and} \quad V_a=\langle\psi_I{}^{(a)}|V_{jl}|\psi_I{}^{(a)}\rangle,$$

the leading terms, of order A and A^2 become for mesons in P-states

$$M_{P-P'}(\pi^+)=-\tfrac{1}{2}K^2A\{\tfrac{1}{2}(V_s+V_a)+\tfrac{1}{16}A(V_s-V_a)\}, \quad (21a)$$

$$M_{P-N'}(\pi^{\pm})=-\tfrac{1}{4}L^2A\{\tfrac{1}{2}(V_s+V_a)+\tfrac{1}{16}A(V_s-V_a)\}, \quad (21b)$$

where $M_{P-P'}(\pi^+)$ refers to π^+ mesons and $M_{P-N'}(\pi^{\pm})$ to π^+ or π^- mesons created in $P-P$ and $P-N$ collisions, respectively. The results obtained for S-state mesons is exactly the same, to this order, and so is that for M''. M''', on the other hand, is really a three-particle problem, and will not be evaluated, since it is expected to be small at 341 Mev. It may be important, however, in explaining the experimentally observed increase of the π^+ to π^- ratio in carbon from 5.1 ± 0.8:1 at 341 Mev[15] to approximately 14:1 at 278 Mev.[16] (The inverse of this ratio was actually measured, by bombarding carbon with neutrons.) A general argument has been advanced by Chew and Steinberger[17] to explain this considerable increase near threshold. The cause is ascribed to the exclusion principle, which should be less important for π^+ production than for π^- production in $P-N$ collisions (i.e., only one particle is excluded on an alpha-particle model in the former case, whereas two are in the latter). It is therefore to be expected that M''' will be larger for π^- than for π^+ production by $P-N$ collisions in nuclei. The energy dependence of the π^+ to π^- ratio, which is due to an effective raising of the threshold for π^- production on the argument of Chew and Steinberger can also be deduced. Near threshold we may expect M''' to be important. At 341 Mev, however, the energy remaining to the struck nucleon in the final state is larger than 80 Mev, as long as mesons of kinetic energy less than 100 Mev are studied. It is thus quite probable that at least one particle will hardly "feel" the exclusion effect and will be ejected from the nucleus. Part of the increase of the π^+ to π^- ratio with decreasing proton energy however may also be accounted for by the difference of the free-particle excitation function and matrix elements involved in the $P-P$ and $P-N$ production processes. Some experimental evidence of this is seen in the fact that Blockman, Passman, and Havens, at Columbia, obtain a ratio of 11 ± 3:1 for carbon bombarded by 381-Mev protons.[5] This increase can hardly be ascribed to an exclusion effect.

At 341 Mev, then, it is expected that either M' or M'' will be of physical interest. If one of these matrices is large, the other one will be small, in general. In order to evaluate either of these matrices explicitly, and hence V_s and V_a, a particular nuclear model must be chosen. We take a rather simple one here, namely a single-particle model, and neglect wave interference.

If the nuclear wave functions are taken as plane waves enclosed in a box in space, of volume $V=(4/3)\pi R^3$ where R is the nuclear radius (Fermi gas model), the matrix M' becomes

$$M'(\pi^+)=-\int[ZK^2+\tfrac{1}{2}(A-Z)L^2]V_K\rho(\mathbf{k})\rho(\mathbf{p}_2)\times\delta(\mathbf{P}+\mathbf{k}-\mathbf{q}-\mathbf{p}_i-\mathbf{p}_2)d\mathbf{k}d\mathbf{p}_2, \quad (22a)$$

[14] E. P. Wigner and E. Feenberg, Reports on Progress in Physics, Phys. Soc., London **8**, 308 (1941).

[15] Richman, Weissbluth, and Wilcox (to be published).

[16] Bradner, O'Connel, and Rankin, Phys. Rev. **79**, 720 (1950).

[17] G. F. Chew and J. L. Steinberger, Phys. Rev. **78**, 497 (1950).

$$M'(\pi^-)=-\int \tfrac{1}{2}(A-Z)L^2V_K\rho(\mathbf{k})\rho(\mathbf{p}_2) \times\delta(\mathbf{P}+\mathbf{k}-\mathbf{q}-\mathbf{p}_i-\mathbf{p}_2)d\mathbf{k}d\mathbf{p}_2, \quad (22b)$$

where $V_K=(4/3)\pi K_{\max}^3$ is the volume of a sphere of maximum momentum, corresponding to the space volume V. The spin sums were evaluated as before, in obtaining Eq. (21), and are correct to the same order. A similar result is obtained for M'', with the roles of $\mathbf{p}_i$ and $\mathbf{p}_2$ interchanged.

$$\rho(k)=\begin{cases}1/V_K & \text{if } 0<k<K_{\max}\\ 0 & \text{if } K_{\max}<k.\end{cases} \quad (23)$$

In the case of a nucleus represented by an excited Fermi gas (i.e., temperature larger than 0°K) we assume that the nuclear wave functions can still be represented by plane waves inside a nuclear volume. In that case (k) is to be interpreted as the momentum of a nucleon on this new model. In this manner the correlation effects can be computed for a general momentum distribution. Equation (22) still holds, but the momentum distribution is now unspecified.

The explicit effect of the exclusion principle is now apparent. If the average value of $\rho(\mathbf{p}_2)$ for the transition process is large, then $M'\sim M$ and the cross section becomes very small. Appropriately, $\langle\rho(\mathbf{p}_2)\rangle_{\text{Av}}$ is large when $\mathbf{p}_2\sim\mathbf{k}$, and both are small in magnitude.

The use of the single particle model in the calculation of the correlation terms, rather than a more general one, neglects the final state interaction of the two colliding particles. Thus, on the Fermi model the cross section does not go to zero, as it should, if the final particles are excluded. In a strictly consistent formulation, both the diagonal and the off-diagonal terms should be treated in the same manner. Hence, on a more general nuclear model than the one used for evaluation of the off-diagonal matrix, a final state interaction term should occur. Since the exclusion terms are expected to be small in any case, we arbitrarily insert a final state interaction here, so that the cross section will at least go to zero correctly on the Fermi model.

If M'' and M''' can be neglected at 341 Mev, the cross section for π^+ and π^- mesons in a nucleus, A, can be written as:

$$\frac{d\sigma_A(\pi^+)}{dTd\Omega_q}=\frac{1}{v_0}\int\left[Z\frac{d\sigma_{P-P}}{dTd\Omega_q}+(A-Z)\frac{d\sigma_{P-N}}{dTd\Omega_q}\right] \times v_R\rho(\mathbf{k})[1-V_K\rho(\mathbf{k}+\boldsymbol{\Delta})]d\mathbf{k}, \quad (24a)$$

$$\frac{d\sigma_A(\pi^-)}{dTd\Omega_q}=\frac{1}{v_0}\int(A-Z)\frac{d\sigma_{P-N}}{dTd\Omega_q} \times v_R\rho(\mathbf{k})[1-V_K\rho(\mathbf{k}+\boldsymbol{\Delta})]d\mathbf{k}, \quad (24b)$$

where $\boldsymbol{\Delta}=\mathbf{P}-\mathbf{q}-\mathbf{p}_i$.

With the inclusion of the correlation effects as treated, it is seen that if L^2 is equal to K^2 the ratio of π^+ to π^- production is still $(Z+A)/(Z-A)$.

E. Absorption and Scattering of Mesons

So far the outgoing meson has been represented by an undamped plane wave. Evidence from photomeson production,[18,19] however, indicates that the meson may interact with other nucleons before it is emitted from the nucleus in which it was produced. The problem of the interaction of mesons with nuclear matter has been treated by Brueckner, Serber, and Watson,[20] who calculated the effect of meson absorption on the production cross section. They quote evidence that this effect is similar for π^+ and π^- mesons, and conclude, furthermore, that the scattering effects are, in general, small compared to the reabsorption ones. They find that the mean free path, λ_a, for meson absorption is of the order of 2–3 a_0, where $a_0=1.4\times10^{-13}$ cm is the effective nucleon radius.

The time taken for a 40-Mev meson to travel a distance of 3 a_0 is of the order of 10^{-22} sec. Since this is small compared to the characteristic nuclear time (taken here as a measure of the time taken for a nuclear disturbance to cross the nucleus, approximately 10^{-22} sec), the probability of meson reabsorption before the nucleus feels the effect of its production is large. Rather than using damped plane waves for the emitted mesons, we therefore treat the meson absorption problem as a separate process.

The information available on absorption, to date, is still scant. For this reason, and due to the other uncertainties involved in the production problem, we neglect the meson energy dependence of the reabsorption process. The analysis of Brueckner, Serber, and Watson shows that the absorption cross section decreases somewhat with meson energy, but is substantially independent of this energy for larger than 40-Mev mesons. On this assumption, and using the model of Fernbach, Serber, and Taylor,[21] they find that the production cross section is reduced by a factor, f_a

$$f_a=3[(1/2x)-(1/x^3)+(1/x^3)(1+x)e^{-x}], \quad (25)$$

where $x\equiv 2R/\lambda_a$, and R is the nuclear radius. For $\lambda_a=3a_0$, and $R=a_0A^{\frac{1}{3}}$, this factor is equal to 0.6.

Experiments performed by Bernardini, Booth, and Lederman at Columbia[22] indicate that, whereas the elastic scattering of mesons by nuclei is indeed small, and approximately energy independent, the inelastic scattering cross section increases rather rapidly with meson energy beyond 30–50 Mev. Thus, the cross section for inelastic scattering at 100–110 Mev is about four times that in the range of 30–50 Mev. The net

[18] R. F. Mozley, Phys. Rev. **80**, 493 (1950).
[19] R. M. Littauer and D. Walker, Phys. Rev. **83**, 206 (A) (1951).
[20] Brueckner, Serber, and Watson, Phys. Rev. **84**, 258 (1951).
[21] Fernbach, Serber, and Taylor, Phys. Rev. **75**, 1352 (1949).
[22] Bernardini, Booth, and Lederman, Phys. Rev. **83**, 1075 (1951).

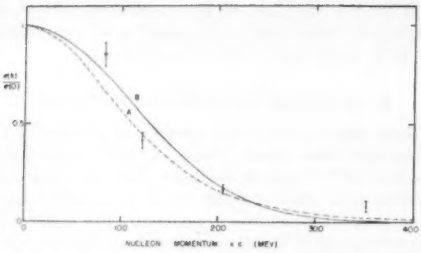


FIG. 1. Modified Chew-Goldberger (A) and Gaussian (B) momentum distributions for a nucleon in C^{12}.

effect of this inelastic scattering is to increase the low energy meson spectrum at the expense of the high energy part. The order of magnitude of this effect can be computed, simply by taking λ_a in x of Eq. (25) to represent the total mean free path for a meson interaction to occur after creation. λ_a decreases then somewhat with meson energy. The effect on f_a is slight, however, for meson energies smaller than 80 Mev. If $f_a=0.6$ at 20 Mev it is only reduced to ~ 0.55 at 80 Mev. The consequences of scattering are therefore neglected in the subsequent calculations.

III. COMPARISON WITH EXPERIMENT

A. Method of Calculation

There is a fair amount of experimental information available at present for meson production in nuclei. The fundamental nucleon-nucleon production data is still rather scarce, however. Whereas the production of positive mesons by $P-P$ collisions has been, and is still being fairly extensively studied both experimentally and theoretically,[1-3,7,23] little is known as yet about the production of either positive or negative mesons in $P-N$ collisions. For this reason, only the experiments of meson production in carbon,[4,15] at 341 Mev are analyzed in terms of the theory developed in the previous section. Both the element and the energy are chosen because of the availability of experimental information.

The equations derived for the meson cross section (both $d\sigma_A^0$ and $d\sigma_A$) involve not only the free transition matrix elements, but also the momentum distribution of the bound, struck nucleon. Theoretically, it is possible to deduce this momentum density from the experimental data. It is simpler, however, to start with a given normalized momentum distribution, such as that of Chew and Goldberger,[24] and to fit it to the experimental cross section. From an analysis of the deuteron pick-up data of York,[25] Chew and Goldberger arrive at the distribution

$$\rho(\mathbf{k})d\mathbf{k}=\frac{\alpha_P}{\pi^2(\alpha_P^2+k^2)^2}d\mathbf{k}, \tag{26a}$$

where α_P corresponds to an energy of 18 Mev. As pointed out in their paper, this momentum density is in doubt at the high end. This distribution, furthermore, gives an infinite average kinetic energy to a nucleon in the nucleus. A preliminary analysis,[26] showed that, due to its long tail, Eq. (26a) did not seem to fit the 90° meson production spectrum obtained by Richman and Wilcox.[4] For this reason, in the calculation made here, the distribution was modified so as to give a finite average kinetic energy of 48.1 Mev, and still fit the experimental points obtained from York's data fairly well. Then

$$\rho_A(\mathbf{k})d\mathbf{k}=C_A\frac{d\mathbf{k}}{(\alpha_P^2+k^2)^2(\beta^2+k^2)^2}, \tag{26b}$$

where C_A is the normalization constant, α_P and β are constants with $\beta=2.5\alpha_P$. This distribution represents the square of the Fourier transform of the wave function $\phi(r)\sim(e^{-\alpha_P r}-e^{-\beta r})/r$. It is shown as curve A in Fig. 1, together with the experimental points of York.

Two other momentum distributions were also investigated. One was a Gaussian type, with an average kinetic energy of 19.3 Mev, chosen so as to still fit the low momentum points obtained by Chew and Goldberger

$$\rho_B(\mathbf{k})d\mathbf{k}=C_B\exp(-\alpha k^2)d\mathbf{k}, \tag{27}$$

where C_B is the normalization constant. ρ_B is shown as curve B in Fig. 1. The third distribution chosen was that for a 0°K Fermi degenerate gas model of the nucleus, with a maximum momentum of 200 Mev/c. A further momentum density has been suggested by Heidmann,[27] namely that of an excited Fermi gas at a temperature corresponding to 9 Mev. This distribution resembles the Gaussian one somewhat, but was not treated, chiefly because of mathematical complications.

Using these momentum densities, assumed to be the same for a proton and a neutron in the nucleus, it is then possible to calculate the energy spectrum at a given angle of emission of the meson. Since the free nucleon production excitation function is unknown, the dependence of K^2, L^2, or of the factors a, b, e, etc., in Eq. (2) on the initial particle momentum is replaced by an average, constant value, determined by fitting the calculated free particle meson spectra to the measured ones at various meson angles. Furthermore, since L^2 is not known we set it equal to ΛK^2 and note whether it is at all possible to fit the experimental carbon data with this choice. It will appear later that this last assumption is not a very good one. This is really not surprising, since

[23] K. M. Watson and C. Richman, Phys. Rev. **83**, 1256 (1951).
[24] G. F. Chew and M. L. Goldberger, Phys. Rev. **77**, 470 (1950).
[25] H. York, Phys. Rev. **75**, 1467 (A) (1949).
[26] E. M. Henley and R. H. Huddlestone, Phys. Rev. **82**, 754 (1951).
[27] J. Heidmann, Phys. Rev. **80**, 171 (1950).

no odd angular dependence can occur in the case of $P-P$ collisions, but may be present in $P-N$ interactions. Thus $c'd'$ in Eq. (2b) need not be zero. Taking $L^2=\Lambda K^2$ implies, in fact, that $\Lambda=a'/a=b'/b=e'/e$ and $c'=d'=0$. Nevertheless, the calculations are made with this choice of L^2, due to lack of more information. The numerical value of Λ is fixed by the π^+ to π^- ratio. In order to fit the observed value of [18] $5.1\pm0.8:1$, $\Lambda=\frac{1}{2}$ is chosen so that the calculated ratio, using either Eq. (17) or Eq. (24) becomes 5:1 (since $d\sigma_{P-N}=\frac{1}{2}\Lambda d\sigma_{P-P}$).

Deuteron formation is neglected in calculating the meson production spectrum in nuclei, since it is expected that the probability of this process is considerably reduced by the nuclear structure. Experimentally, no deuterons have been observed in meson production from carbon.

It follows from the theoretical analysis of W.B. and from the experimental information of Whitehead and Richman[3] that the predominant term in the partial wave analysis of π^+ production in free $P-P$ collisions is one with the meson emitted in a P-state with a $\cos^2\theta$ distribution in the center-of-mass system. In fact, Whitehead and Richman show that the differential cross section in the center-of-mass system of the process as a function of θ, the meson angle is

$$d\sigma_{P-P}/d\Omega=(3.20\pm0.78)(0.071\pm0.068+\cos^2\theta)\times10^{-29}\text{ cm}^2\text{ sterad}^{-1}. \quad (28)$$

If we consider meson S- and P-states only, the observed symmetrical component may be due to either $b(q^2/\mu^2)$ or e (or both) in Eq. (2a). It was shown by Brueckner, Serber, and Watson[28] that some S-state is necessary to explain the experimental results of Panofsky, Aamodt, and Hadley[29] for the capture of π^- mesons in deuterium. They show, in fact, that the ratio of meson S- to P-state [e/a in Eq. (2a)] is approximately $\frac{1}{3}$.

The calculation of typical spectrum—that for π^+ mesons emitted into P-states at 90° to the incident proton beam—is shown in Appendix A. Exclusion effects are not considered in this computation.

The evaluation of the correlation terms can be simplified considerably by choosing an average value of $\mathbf{p}_2$ in Eq. (22). Thus, Passman, Block, and Havens, Jr.[30] assumed that the meson is always emitted with its maximum possible momentum. The justification of this hypothesis rests on the large final state nucleon interaction which occurs when their relative momentum is zero. In this case the magnitude of $\bar{M}$, in general, is very small for mesons emitted both at 90° and at 0° with respect to the incident 341-Mev proton beam. In fact, on the Fermi model, the off-diagonal terms do not contribute at all.

An approximate upper limit to the exclusion effect, on the other hand, is obtained if one of the two final nucleons takes most of the momentum so that little remains to the other one. In this case M'' is certainly small if M' is large and vice versa. The evaluation of M' is carried out in Appendix B, where the approximations employed are also shown. A more thorough treatment of the correlations is not carried out for the present, since the sparseness of fundamental information is not thought to justify it.

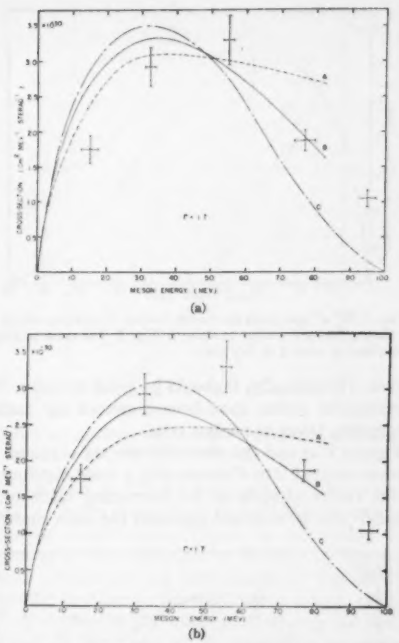


FIG. 2. 90° π^+ spectrum for a meson $P-\cos^2\theta$ state. Exclusion results are neglected in (a) and their approximate maximum effect is shown in (b). Curves A, B, C are for a modified Chew-Goldberger, a Gaussian, and a Fermi distribution, respectively, in this and all subsequent figures.

B. 90° Meson Spectra

On the assumptions discussed in the previous section, the spectra for mesons emitted into P- and S-states at 90° to the incident beam are shown in Figs. 2–4, both without, and with maximum exclusion. In each of these plots, curves A, B, C are the spectra obtained with a modified Chew-Goldberger, a Gaussian, and a Fermi momentum distribution, respectively. The curves are normalized to the free $P-P$ production spectrum, and the absorption factor, $f_a=0.6$ is taken into account. The graphs are multiplied by an arbitrary renormalization constant Γ, chosen so as to obtain an approximate fit to the experimental points, somewhere in between the no-exclusion and the approximate maximum exclusion

[28] Brueckner, Serber, and Watson, Phys. Rev. **81**, 515 (1951).
[29] Panofsky, Aamodt, and Hadley, Phys. Rev. **81**, 565 (1951).
[30] Passman, Block, and Havens, Jr., Phys. Rev. **83**, 167 (1951).

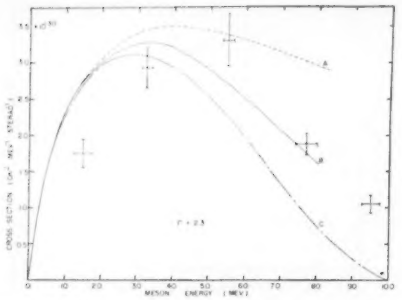


FIG. 3. 90° π^+ spectrum for meson S-state. Exclusion effects are not shown, but are similar to those in Fig. 2. The cross section is normalized to $e/a=1$ in Eq. (2a).

curves. Theoretically, Γ should be equal to unity. The experimental points have been corrected for nuclear absorption, taken as nuclear area.

Figures 2(a) and (b) show the spectra obtained for mesons emitted into P-states with a $\cos^2\theta$ distribution in the center of mass of the interacting nucleons. If $L^2=\Lambda K^2$ this term should represent the main contribution to the cross section. It is noted, however, that $\Gamma=1.7$, so that the calculated spectrum falls considerably too low. It might be remarked at this point that the continuum part of the free $P-P$ cross section, as calculated by W.B. was also too small by 20–30 percent. Not too much emphasis will thus be placed on the values of Γ, though they will be indicated in all instances. Whereas the over-all best fit is obtained with a Gaussian type momentum distribution, it is seen that the calculated spectrum maximum occurs at meson energies which are too small. The calculated spectra also rise too sharply at small meson energies, though the exclusion effect corrects this to a certain degree. At the high energy end of the spectrum, the Fermi momentum distribution cuts the cross section off too rapidly, while the tail of the modified Chew-Goldberger momentum distribution still contributes too much. The Gaussian distribution, on the other hand, follows the spectrum decrease fairly well. At energies beyond those considered, the exclusion effects again lower the calculated spectrum considerably.

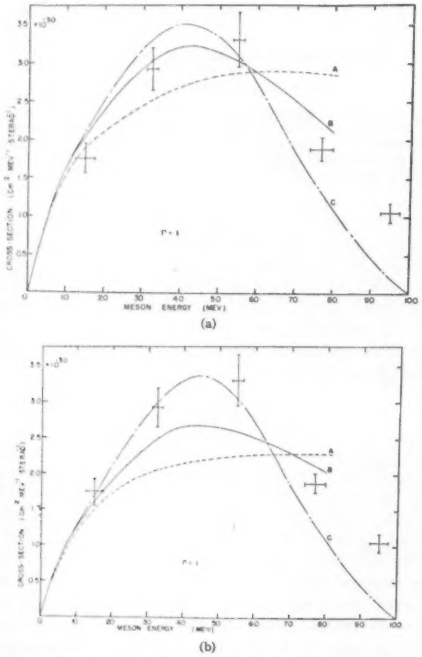


FIG. 4. 90° π^+ spectrum for combined meson $P-\cos^2\theta$, P symmetrical, and S-states, corresponding to $b/a=0.06$ and $e/a=\frac{1}{6}$ in Eq. (2a). Exclusion results are neglected in (a) and their approximate maximum effect is shown in (b).

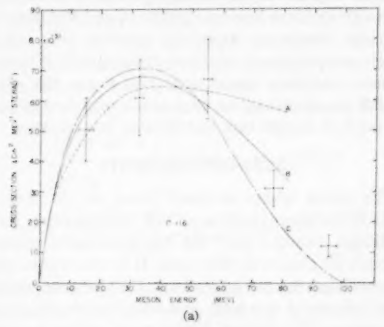
Figure 3 represents the cross section for mesons emitted into S-states, normalized to $e/a=1$ in Eq. (2a). The remarkable fact here is the close correspondence to the $P-\cos^2\theta$ spectrum. In fact, experimentally, except for the difference in Γ, it would be difficult to distinguish between them. For this reason, the combination of $P-\cos^2\theta$ and $\frac{1}{6}$ S-state is not shown. It is obvious, however, that the main effect of the S-state will be in the renormalization constant, Γ, which becomes approximately equal to 1.6.

Figures 4(a) and 4(b) show the spectrum for the case of a combination of meson $P-\cos^2\theta$, symmetrical P, and S-states with $e/a=\frac{1}{6}$ and $b/a=0.06$. The chief result of the 6 percent symmetrical P-state, as expected from its q^2 dependence (this dependence is suppressed by the $\cos^2\theta$ in the $P-\cos^2\theta$ term), is to shift the spectrum peak to higher meson energies, more in agreement with experiment. It also gives a correctly normalized spectrum; that is $\Gamma=1$. On the basis of the 90° experimental spectrum of production, the above combination-type coupling gives the best fit of those considered. It must be remembered, though, that $\cos\theta$ interference terms have not been treated, and these would tend to have a similar, but lesser, effect on the spectrum, as the q^2 term.

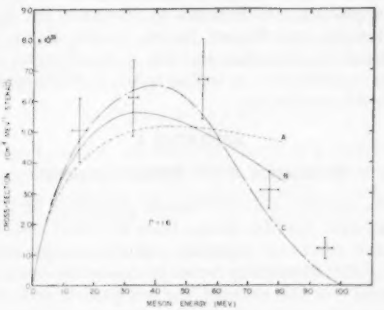
In Fig. 5 the calculated spectrum for combined meson $P-\cos^2\theta$ and S-states with $e/a=\frac{1}{6}$ is compared to the experimental data of Richman and Wilcox (corrected for nuclear absorption) for π^- production at 90°. It is true that the statistical errors are much larger for π^- than for π^+ production, but it is also noted that the experimental spectrum increases more rapidly at low energies, in better agreement with the calculated cross section. Aside from this, no new features appear. The Coulomb effect has been neglected here, as throughout. but it should be small for all but very low energy mesons. The same value of Γ is used as for the equivalent π^+ spectrum.

C. 0° Meson Spectrum

Preliminary experimental results have been obtained by Cartwright,[31] Merrit, Schulz, and Heinz[32] for the production of mesons in carbon, in the direction of an incident 341-Mev proton beam. Their findings can shed considerable light on the matrix element for meson production in $P-N$ collisions. The experimental points together with the calculated spectrum for meson $P\cos^2\theta$ terms is shown in Fig. 6, and for S-state ($e/a=1$) terms in Fig. 7. In each case, the same value of Γ is used as for the corresponding 90° curve. The



(a)



(b)

FIG. 5. 90° π^- spectrum for combined meson $P-\cos^2\theta$ and S-states, corresponding to $b=0$, $e/a=\frac{1}{4}$ in Eq. (2a). Exclusion effects are neglected in (a) and considered in (b).

symmetrical P-wave spectrum was not calculated, but for 0° mesons it should be quite similar to the $P-\cos^2\theta$ one. If $b/a=0.06$, the main effect of the q^2 term will merely be a renormalization of Fig. 6(a) by a factor of about 1.06.

The cross section at 0°, unlike that at 90°, is noticed not to depend greatly on the nucleon momentum distribution. A plausible argument for this result can be

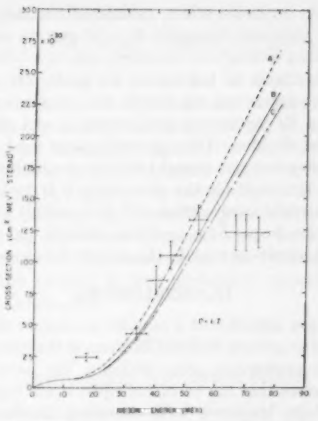


FIG. 6. 0° π^+ spectrum for a meson $P-\cos^2\theta$ state. Exclusion effects are neglected.

made as follows. At 90°, the meson spectrum in nucleon-nucleon collisions is cut off at about 9 Mev by energy-momentum conservation. Hence, in calculating the broadening of this spectrum to 80 Mev by the momentum distribution, the form chosen for this density is expected to influence the spectrum to a large extent. At 0°, however, the free nucleon cutoff does not occur till about 70 Mev, so that in considering the meson spectrum up to 80 Mev, the free-nucleon transition matrix will be of paramount importance.

The $P-\cos^2\theta$ coupling (Fig. 6) cross section is too small at low meson energies, and increases too rapidly at high ones. Both these facts, together with the rapid turning of the experimental spectrum at about 55 Mev, suggest that there is an extensive admixture of S-states and probably of $\cos\theta$ interference terms. Comparison of Figs. 6 and 7 shows that a fairly good fit is obtained if S- and P-states are used in a ratio of 1:1 ($e/a=1$) or more.

The effect of the exclusion terms is not shown in any of these latter figures, but has been calculated. For the

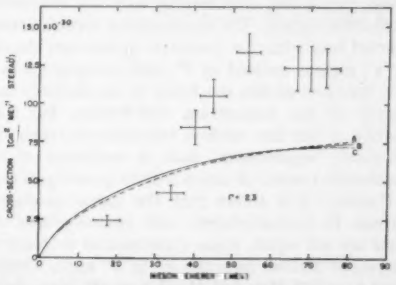


FIG. 7. 0° π^+ spectrum for a meson S-state, normalized to $e/a=1$ in Eq. (2a). Exclusion effects are neglected.

[31] W. F. Cartwright, University of California thesis (April 16, 1951).

[32] Merrit, Schulz, and Heinz (private communication).

maximum exclusion effect, calculated similarly to the example shown in Appendix B, a 50 percent reduction in the cross section was obtained, almost uniformly, so that the shapes of the curves are hardly altered. The computed decrease in the spectra was actually somewhat less than 50 percent at small energies and somewhat higher at 80 Mev. The approximations used in these calculations are not thought strictly justifiable for this large a correction. On the other hand, it is doubted that a more careful computation will give a result which will change the shape of the spectrum greatly, and certainly not sufficiently to explain the experimental one.

IV. CONCLUSIONS

We have seen that it is possible to analyze the meson production process in nuclei in terms of the free nucleon-nucleon production cross sections. We have shown, furthermore that the exclusion effect is not expected to play a very large role in determining the meson spectrum, at the experimental incident proton energies considered (341 Mev), except possibly for mesons emitted in the direction of the beam. Even here, however, these terms will probably not affect the shape of the spectrum to a great extent. The main result of the meson interaction with nuclear matter after creation is an almost uniform reduction of the cross section for the range of meson energies considered. This decrease is chiefly due to meson reabsorption.

The theoretical development was applied to carbon bombarded by 341-Mev protons. Three different momentum distributions were used for a nucleon in C^{12} in an attempt to fit the experimental meson production spectra at 90° and 0°. These were a modified Chew-Goldberger, a Gaussian, and a Fermi distribution with average kinetic energies of 48.1, 19.3, and 12.8 Mev, respectively. The calculated results are still preliminary, due to a lack of knowledge of the fundamental transition matrices. When this information becomes available, it should be possible to decide which of the momentum densities chosen approximates conditions inside the nucleus.

The calculated spectrum at 90° did not depend strongly on the free-nucleon transition matrices involved, but varied considerably with the momentum distribution chosen. The experimental spectrum could be fitted best with the Gaussian momentum density. For π^+ mesons emitted at 0°, with energies up to 80 Mev, the cross section was found to be relatively independent of the momentum distribution, but quite sensitive to the free nucleon transition matrices. The preliminary experimental data is consistent with a considerable amount of meson S-state (possibly as much as P-state). This shows that the meson production matrices in proton-neutron and proton-proton collisions are not equal. Some experimental evidence for this can be derived from the π^+ to π^- ratio, which is 5.1 ± 0.8, at 345 Mev and $11\pm3:1$ at 381 Mev. Initial results obtained for the π^- spectrum at 0° in C^{12} offer further evidence. Its shape seems to be quite different from the π^+ one at the same angle.[22] In this respect, it may be pointed out that, whereas $\cos\theta$ terms (θ is the meson angle in the interaction center-of-mass system) are forbidden by the exclusion effect for meson production in proton-proton collisions, these terms may very well occur in the case of proton-neutron interactions, but have not been considered.

It is thus seen that experiments, presently in progress at Berkeley, on the production of mesons in carbon at 0° are of great interest, as they can shed considerable light on the free proton-neutron meson production cross section. The π^- meson is only produced in such collisions for incident protons. On the other hand, since the exclusion effects are especially small at 90°, and the meson cross section is relatively independent of the free-nucleon transition matrices at this angle, the experimental spectrum can be used as a useful tool to probe the nucleon momentum distribution in a nucleus.

ACKNOWLEDGMENTS

The author wishes to thank above all, Dr. Kenneth Watson for his suggestion of this investigation, for his continued interest and for his invaluable guidance through all phases of this work. It is also a pleasure to acknowledge the aid of Richard Huddlestone with the initial phases of this work and with further stimulating discussions. The author, furthermore, wishes to express his appreciation to Professor C. Richman, Dr. W. F. Cartwright, and Messrs. Merrit, Schulz, Heinz, and Dudziak for discussion and use of experimental data prior to publication, as well as to Mr. C. Wolfe for help with the calculations.

APPENDIX A

Evaluation of 90° Meson Spectrum

The specific case of π^+ meson production at 90° to an incident 341-Mev proton beam is treated here. Correlation effects are neglected, and only mesons emitted into $\cos^2\theta - P$-states by proton-proton collisions are considered. For this example, the free particle production cross section in the center-of-mass system of the interacting nucleons, can be written as:

$$d\sigma_{P-P}^{(c)} = [(2\pi)^4/v_R^{(c)}]\,|(2\pi)^{\frac{1}{2}}\chi_0(p'^{(c)})|^2 \times a(q^{2(c)}/\mu^2)\cos^2\theta\, d\mathbf{J}^{(c)}. \quad \text{(A1)}$$

The superscript (c) is used here to indicate center-of-mass variables. Aside from this, the notation is the same as that used in the main body of the paper. Thus $d\mathbf{J}^{(c)}$ is the phase space volume available

$$\begin{aligned} d\mathbf{J} &= 2\pi M^{\frac{1}{2}}\mu q^{(c)}(T_{\max}{}^{(c)} - fT^{(c)})^{\frac{1}{2}} dT^{(c)} d\Omega_q{}^{(c)}, \\ &= 2\pi M^{\frac{1}{2}}\mu(T_{\max}{}^{(c)} - fT^{(c)})^{\frac{1}{2}} dq^{(c)}/q_0, \end{aligned} \quad \text{(A2)}$$

where $f = 1+\mu/2M$, $T^{(c)}$ is the meson energy, and $T_{\max}{}^{(c)}$ the initial kinetic energy of the nucleons minus

[22] Walter Dudziak, private communication.

μ in the center-of-mass system, computed from Eq. (6), with $B_I - B_F^{(c)}$ neglected. This result also neglects proton-neutron mass differences.

The factor a has been defined in Eq. (2a), and its numerical value is adjusted to fit the experimental spectra at various meson angles. Deuteron formation is taken into account in obtaining the magnitude of a but is neglected in the case of meson production in nuclei.

In order to perform the integration over the momentum distribution of the struck nucleon [see Eq. (17)], the cross section must be transformed to the laboratory frame. This is performed relativistically by making use of the generalized relative velocity definition of Møller,[9] which makes $v_R^{(c)} k_0^{(c)} P_0^{(c)}$ an invariant, where $k_0^{(c)}$, $P_0^{(c)}$ are the initial energies of the interacting nucleons in the center-of-mass system. Since, furthermore $d\mathbf{q}/q_0$ is a relativistic invariant as well as the cross section, we obtain in the laboratory frame

$$d\sigma_{P-P} = \frac{(2\pi)^5}{v_R P_0 k_0} \frac{M^{\frac{1}{2}}}{\mu} a \left\{ P_0^{(c)} k_0^{(c)} |(2\pi)^{\frac{1}{2}} \chi_0(p'^{(c)})|^2 \times \frac{(\mathbf{P}^{(c)} \cdot \mathbf{q}^{(c)})^2}{(P^{(c)})^2} (T_{\max}^{(c)} - fT^{(c)})^{\frac{1}{2}} \right\}_{\text{lab}} \frac{d\mathbf{q}}{q_0}, \quad \text{(A3)}$$

where $p'^{(c)}$ is the relative momentum of the interacting nucleons in the final state and the quantities in the brace are to be expressed in terms of laboratory variables. This can be accomplished relativistically, but it must be kept in mind that $\mathbf{k}$ and $\mathbf{P}$ need not have the same line of action. It thus turns out that the relative velocity of the two systems,

$$\beta = [(P+k_{\parallel})^2 + k_{\perp}^2]^{\frac{1}{2}}/(P_0+k_0), \quad \text{(A4)}$$

where $k_{\parallel}$ and $k_{\perp}$ are the components of $\mathbf{k}$ parallel and perpendicular to $\mathbf{P}$, respectively. Then, referring to Eq. (17) we obtain for the contribution of $P-P$ collisions to the cross section

$$\frac{d\sigma_A^0}{Z} = \frac{d\mathbf{q}}{q_0} \int \frac{(2\pi)^5 a M^{\frac{1}{2}}}{\mu v_0 P_0 k_0} \left\{ \frac{P_0^{(c)} k_0^{(c)}}{(P^{(c)})^2} |(2\pi)^{\frac{1}{2}} \chi_0(p'^{(c)})|^2 \right\}_{\text{lax}} \times \left(U - \frac{\gamma \mathbf{k}}{P_0+k_0} (2\mathbf{P} - f\mathbf{q}) \right)^{\frac{1}{2}} \frac{[\mathbf{q}\cdot\mathbf{k} + q_0(P_0-k_0)]^2}{4} \rho(\mathbf{k}) d\mathbf{k},$$

where $\gamma = (1-\beta^2)^{-\frac{1}{2}}$ and

$$U = 2\gamma[P_0 - P^2/(P_0+k_0)] - 2M + (f-1)\mu - \gamma q_0 f. \quad \text{(A5)}$$

Over the range of integration of k of importance $\beta \sim 0.3$, so that $\gamma \sim 1$. Furthermore, as seen from Eq. (A4), the effect of $k_{\perp}$ on β is quite small, since when $k_{\perp}$ is large, the momentum distribution has cut off, and since $P \gtrsim 2k$ over the integration range. For these two reasons, $k_{\perp}$ is neglected in computing γ. If we take $(2\mathbf{P} - f\mathbf{q})$ to define the z-axis of a cylindrical coordinate system, all but the z-integration can be performed analytically for the momentum distributions chosen. The latter integral is then performed numerically for various values of q_0.

APPENDIX B

Evaluation of Correlation Effects

The effect of the exclusion principle on the cross section, as given by Eqs. (22) and (24), can be calculated on the basis of a single particle nuclear model. The computation involves a double integration over $\mathbf{k}$ and either $\mathbf{p}_2$ or $\mathbf{p}_i$, the other one being taken out by the δ-function in Eq. (22). It was not felt worthwhile carrying out the actual integrals at the present state of experiments. As discussed in the text, an approximate upper limit to the exclusion is obtained if one of the final particles taking part in the reaction has most of the available momentum. In this case, an average value is assumed for $\mathbf{p}_2$ and $\mathbf{p}_i$, consistent with momentum and energy conservation, so that only one integration has to be performed. It suffices to specify the direction of either $\mathbf{p}_2$ or $\mathbf{p}_i$. They are assumed to be along the same line of action. Since $\mathbf{p}_2 + \mathbf{p}_i$ is fixed by momentum conservation, this hypothesis leaves the least momentum to p_i, for a given value of p_2, and the magnitude of these momenta is then entirely determined by momentum and energy conservation. M' can thus be taken to correspond to a small value of p_i, in which case p_2 is sufficiently large that $M'' \approx 0$.

The integration over k is further simplified in this computation by taking $k_{\perp} = 0$ [see Eq. (A4) for the definition of this quantity]. With this assumption, the integration is carried out as that in Appendix A.

Thermal Rayleigh Disk Measurements in Liquid Helium II*

JOHN R. PELLAM AND WILLIAM B. HANSON
Cryogenics Laboratory, National Bureau of Standards, Washington, D. C.

(Received August 1, 1951)

The torque exerted by heat flow on a disk suspended in liquid helium II has been measured and shown to agree with theoretical values. As reported earlier by P. M. Morse and one of the authors, such a disk constitutes a direct mechanical detector of the internal fluid convection currents associated with heat current in helium II. Thus the device is capable of detecting second sound mechanically, whereas ordinary microphones (either pressure or velocity type) do not respond. Furthermore it provides a direct measure of second sound intensity in terms of fundamental quantities, such as kinetic energy density and the geometry. The agreement between observed torque on the disk and predictions based on the two-fluid theory lend further credence to the internal consistency of that hypothesis. Behavior within various temperature ranges, between 1.2°K and the λ-point, strongly implies that both the normal fluid and superfluid components exert their respective contributions to torque on the disk, as if each fluid individually obeyed the classical König formula. That the disk recognizes the net torque exerted by both fluids in a manner conforming to this extension of the original classical hydrodynamics indicates that Bernoulli type forces (and the resultant couples) are associated with each separate fluid component, and with the heat current as a whole. The resultant generalized Bernoulli relationship earlier reported by Morse and one of the authors to hold near the λ-point has now been verified throughout the entire liquid helium II temperature range down to 1.2°K. In view of the generalized form of this principle it would appear that for the majority of classical hydrodynamical devices and experiments there exists the thermal counterpart in liquid helium II. In the present investigation, calculations are based on the numerical values of heat content TS measured for helium II by Kapitza in total flow experiments (through semi-permeable apertures). The present correlations concern alternatively observing by mechanical means the heat current in internal counterflow (no total flow). Employed as a calorimeter the thermal Rayleigh disk yields values of specific heat c_v for liquid helium II in conformity with values obtained earlier by other investigators using customary methods.

I. INTRODUCTION

LIQUID helium II is well known for its ability to transmit thermal waves in the form of self-sustained propagation obeying the wave equation. Although this mechanism is known as second sound, the process is basically thermal in nature, as distinct from ordinary sound. Second sound is generated by means of periodic heating, and is detected by means of the temperature fluctuations accompanying (or constituting) the resultant waves. Ordinary acoustical sources cannot generate such waves nor can ordinary microphones detect them. It is an experimental observation then that second sound is associated neither with net mass flow nor with net pressure fluctuation.

This situation has already been rationalized successfully by both Tisza[1] and Landau[2] on the basis of two-fluid hypotheses for liquid helium II, proposed respectively, by London[3] and Landau.[2] It is supposed that an internal mass flow, or internal convection, between these two fluids replaces the ordinary particle motion of classical sound. Since the oppositely directed momenta of these component fluids exactly balance out, most of the properties of ordinary sound vanish. The further assumption, however, that all heat content of helium II resides in but one of the fluids, normal fluid, leads logically[2,3] to the observed periodic net heat flow and net temperature fluctuations. The manner in which these waves conform to the wave equation, along with their pseudo-mechanical interpretation, gives rise to the terminology[2] "second sound."

Although the two-fluid concept has thus far stood up well, it must nevertheless still be regarded as primarily a useful hypothesis subject to further extensive scrutiny. In this connection the thermal Rayleigh disk experiment[4] provides a new and direct means for observing the presence of internal mass transfer and associated quantities thus postulated for second sound. Not only does the response of this device constitute the first mechanical detection of second sound, but, more important, the quantitative behavior bears directly on the present concepts of the structure of liquid helium II. Thus far, in the temperature range from the λ-point down to 1.2°K, the results are consistent with predictions based on the two-fluid hypotheses, thereby lending credence to this viewpoint. The possibility of measuring certain mechanical and thermal properties of liquid helium II by means of this device at lower temperatures where previous methods fail now appears feasible.

II. THE CLASSICAL RAYLEIGH DISK

The adaptability of the Rayleigh disk to second sound investigations may be demonstrated best by considering first its application to classical acoustics. Extension of the same arguments to the case of thermal waves in liquid helium II then leads systematically to the quantum hydrodynamical case.

The standard classical technique is to expose a small disk to an acoustical field by suspending it at the proper orientation within either a horizontally directed sound beam or a standing sound wave system. A thin torsion fiber attached to the upper edge of the disk holds it in a vertical position within the wave system, the equilibrium orientation of the disk being adjusted by means of the fiber to present an angle of 45° to the propagation vector. In the presence of an ordinary sound field the

* Supported by the ONR.

[1] L. Tisza, J. phys. radium **1**, 165, 350 (1940); Phys. Rev. **72**, 838 (1947).

[2] L. Landau, J. Phys. (U.S.S.R.) **5**, 71 (1941); **8**, 1 (1944).

[3] F. London, Nature **141**, 643 (1938); Phys. Rev. **54**, 947 (1938).

[4] J. Pellam and P. Morse, Phys. Rev. **78**, 474 (1950).

disk tends to swing crosswise to the direction of propagation. Of course when a standing-wave system is involved, the response varies from point to point along the wave pattern; since the mechanism depends upon fluid flow, maximum sensitivity occurs at the position of maximum particle velocity. Our present application involves a completely analogous situation, but with standing thermal waves replacing the standing waves of ordinary sound.

The tendency of a flat obstacle thus to align itself crosswise to the direction of particle excursion in a longitudinal wave was first discussed by Thomson and Tait[5] in 1867. Rayleigh[6] in 1882 described an experiment in which the magnitude of sound waves at the center of a half-wave resonant tube was observed by the deflection of a disk suspended at forty-five degrees to the axis. Based on considerations of the Bernoulli principle, the analytical expression for the torque τ on a disk of radius a suspended in this manner within an incompressible fluid was given by König[7] in 1891 as

$$\tau=(4/3)\rho a^3 v^2 \sin 2\varphi, \tag{1}$$

in terms of fluid density ρ and particle flow velocity v. Since φ is the angle between the disk and the propagation axis, the factor $\sin 2\varphi$ is of course unity for the arrangement used.

For application to cases involving periodic flow the significant feature is the dependence of this expression upon the square of particle velocity v. Accordingly the instantaneous direction of flow is irrelevant and the device functions as a rectifier and thus as a detector; i.e., the disk tends to swing normal to the propagation vector during both portions of the cycle.

The hydrodynamics of this classical situation are represented by the sketch of Fig. 1, adapted from one of Rayleigh's original descriptions.[8] A fluid flow of initial undisturbed velocity v encounters the disk at 45°. Resultant distortion of the flow results in a highly variable velocity distribution over the surface of the disk. For example the splitting of adjacent streamlines on the front side of the disk, at point 1, results in stagnation, whereas at the corresponding point 2 behind the disk a tangential flow velocity persists; similarly, even though stagnation occurs at point 4 where streamlines reconverge behind the disk, flow continues past the front face at point 3.

As a result of these velocity differences between points on the opposite faces of the disk, unbalanced Bernoulli pressures exert a net couple tending to align the disk crosswise to the flow. By way of example, the Bernoulli pressure at point 1 (stagnation) exceeds that at point 2; similarly, the pressure at 4 exceeds that at 3. Accordingly a net couple exists for these sets of

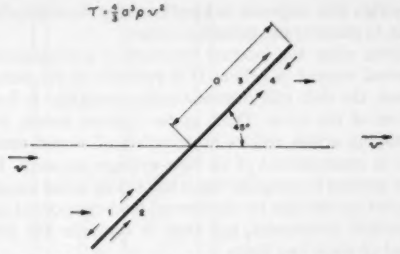


FIG. 1. Fluid flow around a disk. Undisturbed fluid flowing at velocity v.

points, and is exerted in a counter-clockwise direction. The exact treatment by König involves an integration of such a torque (caused by unbalanced Bernoulli pressure) over the entire disk.

For any of the foregoing, an identical procedure for fluid flow in the opposite direction would likewise be valid. The torque would be seen still to persist in the sense tending to align the disk crosswise, hence the detection characteristics of the device in its classical application. In the following section these same principles are extended to embrace the case of internal counterflow about the disk in liquid helium II.

III. THE THERMAL RAYLEIGH DISK

We have just seen that the Rayleigh disk tends to swing crosswise to the propagation axis of an ordinary sound wave regardless of the instantaneous direction (sign) of the particle flow, so that both halves of the cycle contribute constructively. The resultant rectifying action enables this device to detect and measure ordinary sound intensity. This same property of responding to a particle flow irrespective of its actual direction enables the Rayleigh disk to detect the internal counterflow associated with heat flow in liquid helium II.

Even though net particle flow and net momentum remain zero for this internal flow, so that to first order neither vibration microphones nor pressure microphones respond, the disk recognizes variations of Bernoulli-type pressure over its surface. In the same manner that a pressure decrease is associated with regions of increased flow velocity for classical fluids, so does a pressure decrease accompany increased heat current in helium II.

The condition of zero net particle velocity (to first order) for heat flow or thermal vibrations in liquid helium II is customarily written

$$\rho_n v_n+\rho_s v_s=0. \tag{2}$$

Here ρ_n and ρ_s are the densities of the normal and superfluid components, respectively, while v_n and v_s are the corresponding particle velocities. However, although these two individual fluid momenta are oppositely directed, the net energy density is the simple sum of the separate contributions. Otherwise stated, since the

[5] Lord Kelvin and P. Tait, *Natural Philosophy* (Oxford University Press, London, 1867), p. 336.
[6] Lord Rayleigh, Phil. Mag. **14**, 186 (1882).
[7] W. König, Wied. Ann. t. **43**, 51 (1891).
[8] Lord Rayleigh, *The Theory of Sound* (Dover Publications, New York, 1945), Vol. II, p. 44.

Rayleigh disk responds to kinetic energy density rather than to momentum, detection occurs.

Even when the internal counterflow accompanying thermal current in helium II is periodic, as for second sound, the disk still responds indiscriminately to both halves of the cycle. Thus, as for classical sound, this rectifying action results in detection of second sound and in measurement of its time-average intensity. We now proceed to compute separately the expected torques exerted on the disk by the normal fluid component and superfluid component, and then to compute the joint effect of these two fluids.

Assuming that normal fluid interaction with the disk conforms to the classical formula (1) we have for the normal fluid contribution to the torque τ_n when $\varphi=\pi/4$.

$$\tau_n=(4/3)a^3\rho_n\langle v_n^2\rangle_{Av}. \tag{3}$$

Since the mean square value of particle velocity $\langle v^2\rangle_{Av}$ is introduced, Eq. (3) gives the average value of the torque; clearly the disk measures only time-average values. The ac term in the expression for torque contributes no net rotation to the disk, and, in fact,the long time constant of the disk-vane assembly (approximately 10 sec) precludes its response to any of these driving frequencies involved. It is desirable, however, to express v_n directly in terms of heat flow density $\dot{H}$, which may be done by employing the entropy flow relationship[1-3]

$$\dot{H}/T=\rho S v_n, \tag{4}$$

in terms of absolute temperature T, net fluid density ρ, and entropy S (cal g^{-1} deg^{-1}) of liquid helium. Equation (4) merely expresses the (apparently correct) postulate that the entropy of helium II resides entirely within the normal fluid component, and thus that entropy flow $\dot{H}/T$ depends solely upon normal fluid particle velocity v_n. Combining (3) and (4) then, we have

$$\tau_n=(4/3)a^3\rho_n\langle\dot{H}^2\rangle_{Av}/(\rho ST)^2 \tag{5}$$

in terms of $\langle\dot{H}^2\rangle_{Av}$ (time-average value of square of heat current).

It is equally simple to obtain the expression for torque τ_s exerted by superfluid

$$\tau_s=\frac{4}{3}a^3\rho_s v_s^2=\frac{4}{3}a^3\rho_n\left(\frac{\rho_n}{\rho_s}\right)\frac{\langle\dot{H}^2\rangle_{Av}}{(\rho ST)^2}, \tag{6}$$

by using Eqs. (2) and (4) to relate superfluid particle velocity v_s and heat current $\dot{H}$. Here again $\langle\dot{H}^2\rangle_{Av}$ enters to give the time-average value of torque.

Finally, the combined torque exerted by the composite action of the two fluid components on the disk should be the direct sum of these individual contributions. Thus (5) plus (6) results in

$$\tau=\tau_n+\tau_s=\frac{4}{3}a^3\rho\frac{\rho_n}{\rho_s}\frac{\langle\dot{H}^2\rangle_{Av}}{(\rho ST)^2}. \tag{7}$$

Equation (7) could have been obtained alternatively by generalizing the classical Bernoulli relationship† between hydrostatic pressure and fluid flow. By expressing the particle velocities for the individual components in terms of a center-of-mass velocity V plus the relative "thermal velocity" referred to center-of-mass coordinates, the usual kinetic energy term $\frac{1}{2}\rho v^2$ (summed over both fluids) becomes

$$\text{K.E.}=\frac{1}{2}\rho V^2+\frac{1}{2}\frac{\rho_n}{\rho_s}\frac{\langle\dot{H}^2\rangle_{Av}}{(\rho ST)^2}. \tag{8}$$

The second term of (8) applied to the classical Rayleigh disk formula (1) in place of $\frac{1}{2}\rho v^2$ leads to (7).

Actually, considerably more detail should be given in these derivations than has been presented here, and the general situation will be treated more inclusively in a later publication. However, it should be mentioned at this point that the justification for treating the two fluid components separately [as has been done in deriving both Eqs. (5) and (6)] lies in the special isothermal conditions obtaining throughout the central portions of the cavity, and about the disk. Since the disk diameter constitutes but a small fraction of the second sound wavelength involved, additional complications because of temperature gradients are tacitly avoided. For situations involving the observation of pressure differences over distances comparable to a quarter-wavelength (such as the thermal pitot tube)[9] on the other hand, additional considerations are necessary.

The theoretical curves for $\tau_n/\langle\dot{H}^2\rangle_{Av}$, $\tau_s/\langle\dot{H}^2\rangle_{Av}$, and $\tau/\langle\dot{H}^2\rangle_{Av}$ in terms of the preceding expressions (5), (6), and (7) are given as functions of temperature in VI, where comparisons with experimental results are made,

IV. APPARATUS

A composite view of the assembled equipment is presented in the drawing of Fig. 2. The Rayleigh disk assembly itself rests within a Dewar containing liquid helium II. Horizontal slots are used in the Dewar system in addition to the conventional vertical slots, in order to provide apertures for entry and exit of the optical beam employed to observe disk deflections.

The disk D employed for detecting the second sound field consists of a galvanometer mirror $\frac{1}{2}$-inch in diameter. As indicated in Fig. 2 this disk is suspended at the midpoint of a horizontally oriented cylindrical glass cavity E. Resonance of second sound within this cavity is achieved by introducing (electrically) periodic heating at the proper frequency, through the thermal driving element F of carbon resistor strip comprising one end wall of the cavity. The distribution of heat flow density $\dot{H}$ along the axis of the cylinder is indicated for conditions of resonance in the corresponding drawing of Fig. 4.

† *Note added in proof:* While this paper has been in press we have noted that Landau (see reference 2) in his early work mentioned Bernoulli pressures related to individual particle velocities, but without explicit expression in terms of heat current.

[9] J. Pellam, Phys. Rev. **78**, 818 (1950).

Regarding the suspension system, the mirror-disk D is cemented along its vertical diameter to a thin metal rod G supported in turn by the flat (0.0015-in. thick) copper ribbon H. The screw I serves to control both the vertical position and the angular orientation of the mirror by means of rod J extending through a vacuum seal. In the suspended position the metal rod G hangs centrally through a small circular hole L, provided at the upper midposition of E. Above this opening, and thus immersed in helium II but outside the resonance system, damping vanes M composed of thin mica are cemented to G.

A system very similar to that employed for galvanometer suspensions is employed for protecting the torsion ribbon from strain or shock when not in use. By turning I the disk may be lowered when not in use (or prior to transfer of helium) to a position where the lower end of rod G settles into the glass nacelle N. Only after the Dewar system has been filled with liquid helium and cooled to helium II temperatures, thus eliminating all convection currents about vanes M, is the assembly raised to operating position. Adjustment under these conditions to proper height and orientation are then carried out by means of rod J. Finally central adjustment of the rod G within the circular orifice L is accomplished by an over-all vertical adjustment of the entire Dewar assembly.

The small deviations of the disk from its equilibrium orientation caused by the second sound field are observed and measured by the resultant deflection of a light beam. A light source projects light down the axis of the resonant cavity onto the Rayleigh disk D which, being silvered, also fulfills the function of the deflecting mirror. Upon reflection from this disk-mirror the beam emerges at right angles to fall in focus upon a scale (through the horizontal window slots provided for this purpose).

In order to make significant observations on the magnitude of the deflection due to second sound, it is essential that a known and uniform heat flow density be introduced to the resonant system. This is accomplished by means of the electrical heating element F, designed to provide uniform heating over the portion of its surface exposed to cavity E. The thin carbon layer surface of F in contact with helium II supports a uniform electrical current density between two parallel electrode strips R along two opposite edges of the square element (see Fig. 2). This carbon layer actually makes direct contact with helium only over its central circular portion forming the end wall of the cavity; heat generation occurring in the outside "triangular corners" is further suppressed by lacquer applied to these areas.

The driving voltage across F is provided by a signal generator. In order to vary and control frequency within the band-width range (usually less than one cps) additional tuning scales are used. These scales are calibrated by direct (oscillographic) comparison with a standard 100 cps signal, which also serves for frequency stability tests immediately before and after each set of measurements.

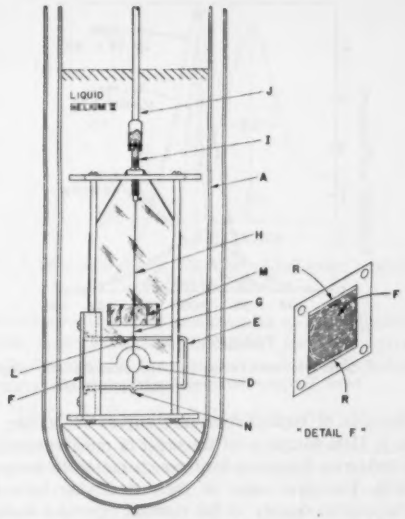


FIG. 2. Over-all drawing showing Rayleigh disk assembly. Detail F shows structure of heater element.

V. MEASUREMENT PROCEDURES

(a) Resonance Response

The virtually complete inherent quiescence of helium in the liquid II state contributes significantly to the stability of the system, and, accordingly, to the reproducibility of the data obtained. In particular the property of temperature uniformity entirely precludes the formation or presence of any convection currents. As a result, the disk-vane system is suspended under such stable conditions that the total zero point shift of the indicator spot on the scale 50 centimeters from the disk can be less than one-half millimeter during the course of several hours.[10] Only within the resonance chamber E, in the presence of second sound waves sustained by an externally maintained driving stimulus, does the suspended system experience a torque.

Ambient temperature is maintained sufficiently constant for measurements of this nature by setting the pumping rate constant and correcting for subsequent small fluctuations by means of a separate variable light source directed against the helium bath; such temperature fluctuations as do occur are observed (for correction) in terms of helium vapor pressure variations by means of an absolute oil manometer read through a telescope.

[10] On the other hand for temperatures above the λ-point the turbulent helium I may throw the indicator spot completely off scale.

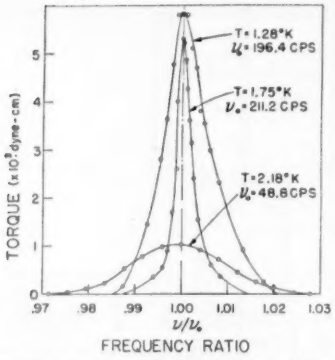


FIG. 3. Typical resonance curves for three characteristic temperatures. Torque τ *versus* relative frequency ν/ν_0.

Samples of typical resonance curves are shown in Fig. 3. Here torque τ (of the order of millidyne-cm) is plotted *versus* frequency for three characteristic temperatures. The same value of thermal driving force $\dot{H}_0$ (heat current density at the thermal generator surface F) was maintained for each case, and, in order further to compare the resultant resonance response, torque τ has been plotted *versus* relative frequency ν/ν_0. That is, the system must be retuned for each temperature (because of temperature-dependency of wave velocity) so that resonant frequency ν_0 is different for each response curve (as indicated, Fig. 3). In this regard, data on wave velocity *versus* temperature provided by the Rayleigh disk displays the usual characteristics, i.e., the maximum of (about) 20.3 m/sec at 1.65°K falling off somewhat at lower temperatures and dropping abruptly toward zero with temperature increase toward the λ-point.

As evident from these curves, both the shape and magnitude of the resonance response (for constant heat input $\dot{H}_0$) vary with temperature. Although discussion of the physical significance of these tendencies will be deferred until later, we point out here that two essential criteria are provided by this data. First, the maximum torque $\tau_{\max}$ exerted by the second sound field on the disk may be observed directly from the resonance peak. Second, the amount by which the basic heat current density $\dot{H}_0$ applied to the cavity is augmented by resonance reenforcement for these peak conditions is revealed by the band width (sharpness) of the response.

In close analogy to the treatment for classical resonator systems, a re-enforcement factor f for the present experiment may be represented by

$$f=\dot{H}/\dot{H}_0 = (2/\pi)(\nu_0/\Delta\nu)=(2/\pi)Q, \tag{9}$$

where $\Delta\nu$ is a band width. More specifically, $\Delta\nu$ denotes the frequency spread between the two points on the response curve (on either side of resonance) for which torque τ drops to one-half maximum value. The quantity f is the factor by which heat current input $\dot{H}_0$ at the generator wall is "amplified" by resonance to the effective value $\dot{H}$ for the mid-portions of the cavity. The quantity $\nu_0/\Delta\nu$ is, of course, just the usual Q for a resonant system. Relationship (9) will be derived in a following section.

(b) Second Sound Generation

Second sound waves are generated in the cavity E at twice the frequency of the periodic voltage applied across F. Thus, if the effective voltage is V_{rms} at frequency $\nu/2$, the instantaneous heat current injected into the resonator becomes

$$\dot{H}_0=(1/4.18A)(2V_{\text{rms}}^2/R)\cos^2 2\pi(\nu/2)t, \tag{10a}$$

$$=(1/4.18A)(V_{\text{rms}}^2/R)[1+\cos 2\pi\nu t]. \tag{10b}$$

The area A appears in the denominator since $\dot{H}_0$ is heat flow per unit area, and the additional numerical factor converts the units to cal sec^{-1} cm^{-2}; resistance R between the electrodes of F is measured for each set of measurements. The first term of (10b) represents a constant, or background, heat input density while the second is the actual ac heating term generating second sound at frequency ν. The periodic component is selectively enhanced by resonance within the cavity.

As we have already seen, the torque exerted against the disk varies as heat current squared, so that the actual quantity in which we are interested is the average value of

$$(\dot{H})^2=(1/4.18A)^2(V_{\text{rms}}^4/R^2)[1+f\cos 2\pi\nu t]^2 \tag{11}$$

the resonant reinforcement factor f (Eq. (9)) being introduced only for the periodic component. The dc heating term does not gain from resonance, and furthermore becomes dissassociated with the standing wave system by promptly escaping through the exit orifice L.

Accordingly the effective value of $\dot{H}^2$ becomes merely ($f\gg 1$)

$$(\dot{H})^2=f^2(1/4.18A)^2(V_{\text{rms}}^4/R^2)\cos^2 2\pi\nu t, \tag{12}$$

with average value

$$\langle\dot{H}^2\rangle_{\text{Av}}=\tfrac{1}{2}f^2(1/4.18A)^2(V_{\text{rms}}^4/R^2). \tag{13}$$

Here $\dot{H}$ is the thermal counterpart of the undisturbed flow velocity v (Eq. (1) and Fig. 1) for the classical application of the Rayleigh disk. It is the effective heat flow density developed near the midportions of the cavity.

(c) Second Sound Field and Band Width

The analysis of such a resonant system as considered here may of course be carried out by standard techniques for treating standing wave situations. Although such approaches are perfectly adequate and more direct, many extraneous questions which arise during their application to this particular case may be tacitly

avoided[11] by considering directly the multiply reflected waves within the cavity. The latter method will accordingly be employed.

The thermal generator F has been seen to emit a periodic heat current of frequency ν so that a continuous travelling second sound wave represented by $H_0 e^{2\pi\nu i(t-x/v_2)}$ is introduced to the resonator. Here v_2 is the wave velocity and x the distance from F measured along the axis of E as indicated in Fig. 4. Restricting our considerations now to the value of the heat current $\dot{H}$ at the midplane of the cavity ($x=L/2$), the wave system cumulates after successive reflections to

$$\dot{H}=\dot{H}_0 e^{2\pi\nu i(t-x/v_2)}[1-re^{-2\pi\nu iL/v_2}+r^2e^{-2\pi\nu i(2L/v_2)} -r^3e^{-2\pi\nu i(3L/v_2)}+\cdots]. \quad (14)$$

Here the first term represents the initially introduced wave,[12] with successive terms accounting for ensuing multiple reflections. The boundary conditions of (nearly) zero heat conductivity at either endwall of the cavity require the alternate reversals in sign. The factor r represents the fraction of the wave persisting between any two successive transits past the center point ($x=L/2$). The persistence factor r defined in this manner is very near to unity since the system is highly resonant; however r tacitly takes into account losses from any causes such as reflection from the endwalls, attenuation within the liquid itself, and dissipation about the disk. Since the system has a high Q the simplifying assumption of a single r is justified.

Numerical evaluation of the resonance properties of the system follow directly from band-width observations. In this connection expression (14) may be written more conveniently as

$$\dot{H}=\dot{H}_0 e^{2\pi\nu i(t-L/2v_2)}/[1+|r|e^{-2\pi\nu i(L/v_2)+i\theta}], \quad (15)$$

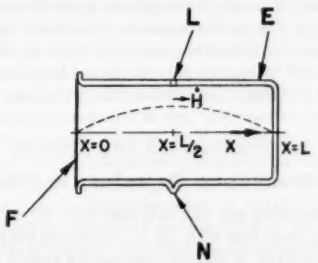


FIG. 4. Thermal half-wave resonator. Heat flow distribution within resonant cavity (dotted line $\dot{H}$).

[11] The standing wave analysis would require setting up an expression for impedance (i.e., "thermal impedance") as a function of distance x along the cavity axis. Damping responsible for the frequency band width would then have to be associated arbitrarily either with losses (attenuation) within the liquid, and/or losses at the tubewalls and endwalls, and/or losses occurring about the disk (particularly at the edges). These unnecessary considerations are automatically avoided in the present analysis.

[12] The known temperature discontinuity existing at the interface between a solid and liquid helium II in the presence of heat current is of no consequence to the present treatment. The temperature fluctuation of surface F plays no role in the analysis, since we are concerned only with the amount of heat originating at the thermal generator.

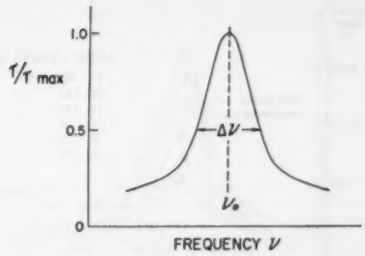


FIG. 5. Bandwidth criteria. Definition of half-torque bandwidth $\Delta\nu$ [see Eq. (18)].

where now r is treated as a complex quantity (possible phase angle $\theta\ll\pi$). The measured torque τ exerted upon the disk is then proportional to $\dot{H}^2$, so that observation of the quantity

$$|\dot{H}^2|=\dot{H}_0^2/[1+2|r|\cos(2\pi\nu L/v_2+\theta)+|r|^2] \quad (16)$$

[derivable from (15)] leads to these determinations. As a result of the quadratic dependence of torque on heat current, band width is measured between frequencies giving half-maximum torque as shown in Fig. 5 (in contrast to linear systems, for which the $1/\sqrt{2}$ value is used).

The resonant frequency ν_0 occurs for $2\pi\nu_0 L/v_2+\theta=\pi$, so that in terms of the amount $(\nu-\nu_0)$ off-resonance the torque becomes proportional to

$$|\dot{H}^2|=\dot{H}_0^2/\{1+r^2-2r\cos[\pi(\nu-\nu_0)/\nu_0]\} \simeq\dot{H}_0^2/\{(1-r)^2+[\pi(\nu-\nu_0)/\nu_0]^2\}, \quad (17a)$$

the latter approximate form holding for $r\simeq 1$ and $[(\nu-\nu_0)/\nu_0]\ll 1$. Clearly half-maximum torque occurs for the condition $\pi(\nu-\nu_0)/\nu_0=\frac{1}{2}\pi\Delta\nu/\nu_0=(1-r)$, which determines r in terms of band width.

Augmentation of torque τ becomes accordingly

$$|\dot{H}^2|/|\dot{H}_0^2|=\{(\tfrac{1}{2}\pi\Delta\nu/\nu_0)^2+[\pi(\nu-\nu_0)/\nu_0]^2\}^{-1}, \quad (17b)$$

so that for directly on resonance (τ_{max})

$$|\dot{H}^2|/|\dot{H}_0^2|=f^2=(2/\pi)^2(\nu_0/\Delta\nu)^2 \quad (17c)$$

verifying the expression for f previously given, Eq. (9). Or, in terms of torque ratios, we have (see Fig. 5)

$$\tau/\tau_{max}=\{1+[2(\nu-\nu_0)/\Delta\nu]^2\}^{-1} \quad (18)$$

for our resonance response.

(d) Distortion of Heat Current by the Disk

During the preceding derivation no account has been taken of the distortion of the wave at the midpoint of the cavity by the presence of the disk. Actually this is justifiable, since the expressions for torque on the Rayleigh disk are consistently in terms of the unperturbed fluid velocity (or heat current) in the absence of the disk. The König formula (1) itself takes into account the distortions set up by the disk as an obstacle; and the fact that, for the present application, each portion of the wave train passes the disk a large number

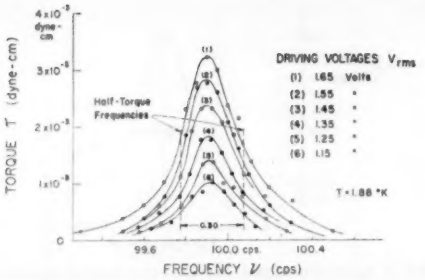


FIG. 6. Band width independent of power level. Series of resonance curves (τ *versus* ν) at constant ambient temperature T (1.88°K) for various voltages V_{rms}; half-torque frequencies are indicated by parallel broken vertical lines (constant band width $\Delta\nu = 0.30$ cps).

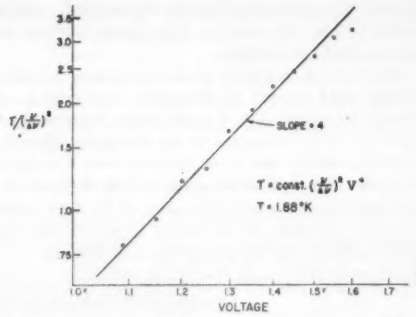


FIG. 7. Fourth power dependence of torque on voltage. Measured $\tau_{max}/(\Delta\nu/\nu_0)^2$ *versus* driving voltage V_{rms}; slope equal to four on log-log scale ($T = 1.88$°K).

of times before dying out does not change the situation. It is important only to realize that the net losses incurred by the waves (and measured by observing $\Delta\nu$) are for the system as a whole, and may include dissipative effects about the disk.

The only way in which the distortion of the wave pattern by the disk could affect the results would be in terms of interactions between the suspended disk and the walls of the resonant cavity. Although such interaction could become important for disk diameters nearly as large as the tube diameter, this is not the case for the present situation. For example, carefully conducted experiments on the classical Rayleigh disk in air by an earlier investigator[13] have shown that until disk diameters exceed roughly one-half of the tube diameter, corrections remain only a few percent (i.e., less than 4 percent). It is reasonable to suppose that the same criteria should hold for the corresponding situation involving thermal waves, and accordingly no corrections have been included for such effects.[14]

Similarly the ratio of disk diameter to wavelength[15] remains well below the criterion value[16] of 0.2 for legitimate application of the König expression.

(e) Defects of Classical Rayleigh Disk Absent

Actually the present use of the Rayleigh disk for detection of thermal waves in liquid helium II represents a far more ideal situation than had previously been encountered with classical applications of this device. Not only is virtually complete mechanical stability attained (quiescence of helium II and complete absence of convection currents) but also various first-order corrections are inherently eliminated by the nature of the internal counter-flow involved. Primary among these are the effects of finite mass of the disk, and resonance in flexural vibrations of the disk.

For example in classical sound waves the first-order pressure fluctuations and mass flow result in forced oscillations (longitudinal, along the propagation axis) of the disk as a whole; as demonstrated theoretically by King,[17] the associated reduction in particle flow velocity relative to the disk decreases the observed torque. This may become serious, particularly for acoustic waves in liquids, where measured reduction factors as high as 2500-fold have actually been reported.[18] It is noteworthy, on the other hand, that for thermal waves in helium II such defects are entirely absent. The inherent nature of the propagation, including complete absence of first-order pressure fluctuations and mass flow, precludes such forced oscillations of the disk, thus eliminating finite mass effects. Correspondingly, for similar reasons, anomalous defects in the classical application caused by flexural vibrations of the disk[16] (occurring only at certain frequencies) are likewise avoided, there being no first-order driving forces.

VI. OBSERVATIONS AND RESULTS

(a) Fourth Power Dependence on Voltage

By combining any of the torque Eqs. (5), (6), or (7) involving heat flow with Eq. (13) relating heat flow to applied voltage, it follows that torque should vary as the fourth power of voltage. Specifically, for the composite effect of both fluids, the expression for net torque τ (τ_s plus τ_n) becomes

$$\tau = \frac{8}{3\pi^2}\left(\frac{1}{4.18A}\right)^2 \frac{\rho a^3}{(\rho S T)^2}\left(\frac{\rho_n}{\rho_s}\right)\left(\frac{\nu_0}{\Delta\nu}\right)^2\left(\frac{V_{rms}^4}{R^2}\right). \quad (19)$$

Clearly, it is necessary to establish this relationship experimentally before proceeding with confidence to the general application of such expressions. Also ques-

[13] R. A. Scott, Proc. Roy. Soc. (London) **A183**, 296 (1945).
[14] For the present investigation the ratio of disk to tube diameter was just one-half; i.e., disk diameter ½ in., tube diameter 1 in. Even if corrections were important, they would involve but a constant scale-factor, independent of wave velocity (and T).
[15] Disk diameter = 1.27 cm; wavelength $\lambda = 9.50$ cm.
[16] E. Barnes and W. West, J. Inst. Elec. Engr. (London) **65**, 871 (1927).
[17] L. King, Proc. Roy. Soc. (London) **A153**, 17 (1935).
[18] A. Wood, Proc. Phys. Soc. (London) **47**, 779 (1935).

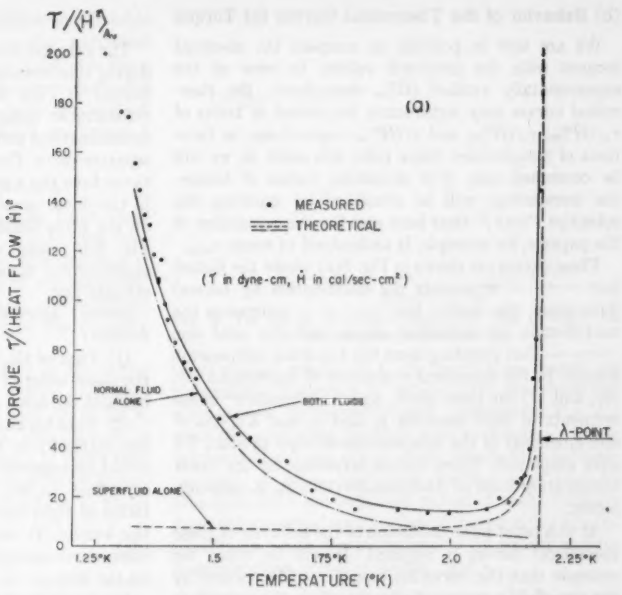


FIG. 8(a). Torque ratios $\tau/\langle\dot{H}^2\rangle_{hr}$ as functions of temperature T: experimental and theoretical values. Theoretical predictions indicated by curves: — · — · — represents $\tau_n/\langle\dot{H}^2\rangle_{hr}$ the expected torque for normal fluid; — — — — represents $\tau_s/\langle\dot{H}^2\rangle_{hr}$, the expected torque for superfluid; and ——— represents $\tau/\langle\dot{H}^2\rangle_{hr}$ that predicted for both. The circles are the observed values.

tions of linearity (i.e., independence of band width on power level) need be investigated.

In this connection, a representative family of resonance curves is given in Fig. 6 for an arbitrary ambient bath temperature T of 1.88°K. These curves‡ of measured torque τ *versus* frequency ν were taken for a set of different driving voltages V_{rms}, as indicated. Not only is the sensitive dependence of torque τ on voltage evident from Fig. 6, but also the comparative independence of band width $\Delta\nu$ on voltage. Thus the two parallel vertical lines (a) and (b), separated by frequency difference $\Delta\nu = 0.30$ cps, are seen to intercept each curve at very near to its half-torque values (or half-power values, torque being proportional to energy density). This is taken to indicate that for this range of torque (i.e., for this magnitude of "strain" imposed upon the liquid helium II in maintaining the torque) the system is "linear" and that no complicated higher order effects are contributing.

As to the precise dependence of maximum torque on voltage, Fig. 7 shows a quantity involving τ_{max} plotted *versus* V_{rms} on a log-log scale. For generality [to conform to Eq. (19)] the quantity $\tau_{max}/(\nu_0/\Delta\nu)^3$ is actually plotted; this amounts, however, to a mere additional factor, because of the virtual constancy of $\nu_0/\Delta\nu$. That the fourth power dependence on voltage holds is amply demonstrated by the slope 4.0 for this set of points. Accordingly Eq. (19), and the two associated forms pertaining to the individual fluid components separately, may be employed confidently in so far as the voltage dependence is concerned, within this range of measured torque. Dependence on the remaining parameters of (19) will be examined shortly.

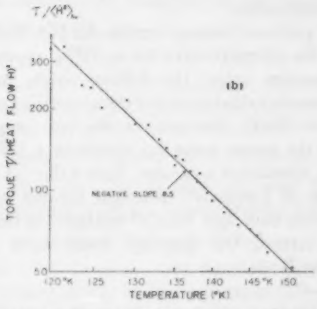


FIG. 8b. Log-log plot of $\tau/\langle\dot{H}^2\rangle_{hr}$, *versus* T. Circles represent measured values. The solid line drawn through these points has a negative slope of 8.5.

‡ *Note added in proof:* A slight asymmetry may be noted in these resonance curves. This is a higher order effect resulting from the slight dependence upon disk orientation of the resonance frequency of the disk-cavity system. That is, deflections of the disk from equilibrium lengthen the effective second sound path about the obstacle slightly, and accordingly lower the resonance frequency. Thus the peaks of Fig. 6 are shifted progressively more and more to the left with increasing height, as may be seen. The measured results (band width, or τ_{max}) are in no way affected by this action, which in fact could be eliminated by using small enough deflections (or a smaller disk).

(b) Behavior of the Theoretical Curves for Torque

We are now in position to compare the observed torques with the predicted values. In view of the experimentally verified $\langle \dot{H}^2 \rangle_{Av}$ dependence, the theoretical curves may legitimately be plotted in terms of $\tau_n/\langle \dot{H}^2 \rangle_{Av}$, $\tau_s/\langle \dot{H}^2 \rangle_{Av}$, and $\tau/\langle \dot{H}^2 \rangle_{Av}$, respectively, as functions of temperature. Since from this point on we will be concerned only with maximum values of torque, the terminology will be simplified by omitting the subscript "max;" thus here and for the remainder of the paper τ, for example, is understood to mean τ_{max}.

These curves are shown in Fig. 8(a) where the dotted line —·—·— represents the contribution by normal fluid alone, the dashed line — — — — represents the contribution by superfluid alone, and the solid one ———— that resulting from the two fluid components jointly. In the numerical evaluation of expressions (5), (6), and (7) for these plots, Andronikashvilli's[19] measurements of fluid densities ρ_n and ρ_s and Kapitza's[20] measurements of the temperature-entropy product TS were employed. These curves terminate at the lower temperature limit of Andronikashvilli's ρ_n/ρ_s measurements.

At this point some discussion of the behavior of these theoretical curves is required. It will be noted for example that the curve for torque $\tau_s/\langle \dot{H}^2 \rangle_{Av}$ caused by the superfluid component rises sharply in the immediate vicinity of the λ-point. Thus the superfluid contribution appears to be enhanced most in the temperature range for which the superfluid concentration is least. This takes place because as superfluid becomes sparse its particle velocity v_s increases to maintain zero net momentum transfer [according to Eq. (2)]; and the simultaneous dependence of torque $\tau_s/\langle \dot{H}^2 \rangle_{Av}$ on particle velocity squared more than offsets the effects of decreased density. It may likewise be observed from the curve representing $\tau_n/\langle \dot{H}^2 \rangle_{Av}$ that at the low temperature limit the torque caused by the normal fluid component in turn increases many fold, and for similar reasons; in that temperature range the sparseness of normal fluid is responsible.

Other points of interest include the fact that at the λ-point the (dotted) curve for $\tau_n/\langle \dot{H}^2 \rangle_{Av}$ approaches a finite, nonzero value; this follows simply from the finite, nonzero value for each of the quantities appearing in (6) as $T \to T_\lambda$. Similarly at the low temperature extreme the torque curve for superfluid $\tau_s/\langle \dot{H}^2 \rangle_{Av}$ presumably approaches a nonzero, finite value (at least on the basis of Landau's[21] treatment for the very low temperature limit; had Tisza's[22] treatment in this region proved correct, the quantity would have become infinite as $T \to 0$).

[19] E. Andronikashvilli, J. Exp. Theor. Phys. (U.S.S.R.) **18**, 424 (1948).
[20] P. Kapitza, J. Phys. (U.S.S.R.) **5**, 59 (1941).
[21] L. Landau, Phys. Rev. **75**, 884 (1949).
[22] L. Tisza, Phys. Rev. **75**, 885 (1949).

(c) Measured Torque

The values of torque ratio $\tau/\langle \dot{H}^2 \rangle_{Av}$ actually measured during this investigation are indicated by the circles • plotted in Figs. 8(a) and (b). To accommodate the considerable range of magnitude over which these determinations extend, the results are presented in two separate plots. Thus measurements in the temperature range from the λ-point down to about 1.3°K are shown in the linear graph of Fig. 8(a). Results below 1.3°K on the other hand are included in the log-log plot of Fig. 8(b); here no theoretical curve is given, the experimental slope being indicated instead by the straight line.

Several aspects of these results are immediately evident:

(1) First of all, not only are the observed torques of the same order of magnitude as the predicted values, but also the same general trends are observed.

(2) Thus for example the measured values of $\tau/\langle \dot{H}^2 \rangle_{Av}$ rise markedly in the neighborhood of the λ-point, as would be expected on the basis of the contribution by superfluid. In fact, $\tau/\langle \dot{H}^2 \rangle_{Av}$ is observed to increase by a factor of eight between the temperatures of 2.1°K and the λ-point. It would thus appear that the superfluid component contributes its share to the torque exerted on the disk.

(3) Similarly for temperatures below about 1.75°K (and down as far as measurements have thus far been carried, 1.2°K), the torque $\tau/\langle \dot{H}^2 \rangle_{Av}$ likewise displays a marked increase. This is in accordance with the expected behavior caused by normal fluid. Accordingly the normal fluid component apparently contributes its share to the torque exerted on the disk.

(4) In the intermediate temperature range (from about 1.75°K to about 2.1°K) where neither fluid component is expected to predominate, the observed results conform closely to the expected combined effects of the individual fluid components. Since in fact, over the entire temperature range thus far investigated the observed torques are in conformity with the composite (solid) curve predicted for the combined effects of both fluids, it would thus appear that each fluid component contributes its share to the torque on the Rayleigh disk, and that the total observed torque is the sum of these individual torques. This is interpreted as fully verifying Eq. (7).

These same observations may otherwise be specified by stating that torque τ on the disk is determined by the total (average) kinetic energy density (resulting from normal fluid plus superfluid), associated with the internal counterflow. On this basis the results presented in Fig. 8 provide verification of the generalized Bernoulli relationship

$$\frac{1}{2}\rho V^2+\frac{1}{2}\rho\frac{\rho_n}{\rho_s}\frac{\langle \dot{H}^2 \rangle_{Av}}{(\rho ST)^2}+\rho gh+p=\text{constant} \qquad (20)$$

given by Morse and one of the authors[4,9] during the early stages of this investigation. Here the second term

is the contribution to hydrostatic pressure attributable to heat current, and is identical to the additional kinetic energy density term appearing in Eq. (8). Since as already stated [after (8)], this factor replaces the usual kinetic energy term for the classical Rayleigh disk formula (1), the verification of (7) is likewise a verification of (20).

(d) Measurement of Q ($Q=\nu_0/\Delta\nu$)

Data of secondary interest resulting from these investigations include measurement of Q for the configuration employed. As stated in (9) this quantity Q retains its usual definition, the ratio of resonance frequency ν_0 to half-power (or half-torque) band width $\Delta\nu$ so that $Q=(\pi/2)f$. Figure 9 is a plot of this quantity *versus* temperature.

It is apparent from this curve that in the upper temperature range the Q for the system decreases markedly from a maximum value of about 340 at 2.0°K toward (probably) zero at the λ-point. This behavior is consistent with expectations, since ideally Q would bear a relationship to the temperature attenuation coefficient α for second sound. That is α (in units of cm^{-1}) would be given by

$$\alpha=(2/\pi)(1/QL), \tag{21}$$

if losses were attributable solely to second sound attenuation; here L is the length (4.75 cm for the present case) of the resonant cavity. And previously existent qualitative measurements of α have indicated increased attenuation[23] near the λ-point. It is emphasized however that (21) is not applicable directly to the present case because it concerns only losses attributable solely to second sound attenuation. Since losses may well occur at the resonator walls and especially around the disk (presumably at the edges), the observed variation of Q with temperature is probably far more complicated than implied by (21). This should be the case particularly at the lowest temperatures where increased normal fluid velocities may introduce turbulence about the disk.

Finally the temperature dependence of maximum deflection is of some interest, explaining for example the trends earlier noted in the response curves of Fig. 3. Values of τ_{max} plotted *versus* temperature, at constant power input level $\dot{H}_0$, would show a rise in observed torque at the lowest temperatures in spite of the increases losses in that range. On the other hand, the rapid increase in losses near the λ-point more than offsets the intrinsic increase in $\tau/\langle\dot{H}^2\rangle_{Av}$, so that the torque τ (as distinct from $\tau/\langle\dot{H}^2\rangle_{Av}$) is actually observed to fall off in that range.

VII. THERMODYNAMIC INTERPRETATIONS

Although the original expectation[4] of this investigation was to produce an observable effect, the precision of measurement now attained allows independent investigation of some of the thermodynamic properties of liquid helium II. Thus the correlation between our

[23] J. Pellam, Phys. Rev. **75**, 1183 (1949).

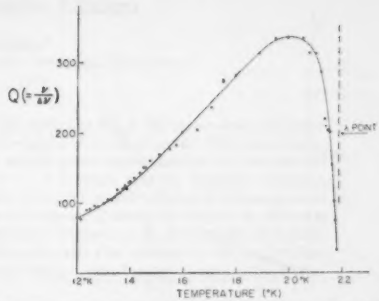


FIG. 9. The cavity Q *versus* temperature. The $Q(=\frac{1}{2}\pi f)$ for second sound resonance in cavity.

experimental results and the predictions based on expression (7) appears to substantiate the numerical evaluation of the thermodynamic combination $S^2\rho_s/\rho_n$; since this evaluation depends both on Kapitza's measurements of entropy S and Andronikashvilli's determinations[19] of ρ_n/ρ_s, additional credence may thus be associated with the published values of each. Furthermore the numerical specification of this quantity $S^2\rho_s/\rho_n$ constitutes a new thermodynamical constraint leading to independent values of specific heat c_v.

(a) The Thermal Rayleigh Disk as a Calorimeter

It will be within the limits of this investigation to consider briefly such calorimetric applications of this device. The preceding treatment has been based solely on the two-fluid hypothesis, plus Landau's (apparently correct) further assumption that all thermal content ST resides exclusively in the normal fluid component. The analysis has concerned primarily a resonant wave system, without particular emphasis on the thermodynamic considerations. However, the preceding thermodynamic combination $S^2\rho_s/\rho_n$ may alternatively be lumped in terms of an independently measurable quantity, the wave velocity v_2 of second sound.

The most convenient expression for v_2 for this purpose is the modification by Peshkov[24] of the Tisza-Landau formulation, namely

$$v_2^2=(\rho_s/\rho_n)(TS)^2/c_v.$$

The generalized Bernoulli expression formulated in these terms has been given earlier (prior to these measurements) by one of the present authors[9] as

$$\tfrac{1}{2}\rho V^2+\rho gh+p+\tfrac{1}{2}(1/\rho c_v T)(\dot{H}/v_2)^2=\text{constant}. \tag{22}$$

On this basis the corresponding expression for total torque on the disk reduces to a form equally valid to (7)

$$\tau=(4a^3/3)(1/\rho c_v T)(\dot{H}/v_2)^2, \tag{23}$$

and from which we could plot a function equivalent to the solid curve of Fig. 8. This expression (23) actually can be obtained by quite independent methods, as will be shown shortly by one of the present authors, and thus may be regarded as just as basic as expression (7).

[24] V. Peshkov, J. Exp. Theor. Phys. (U.S.S.R.) **18**, 857 (1948).

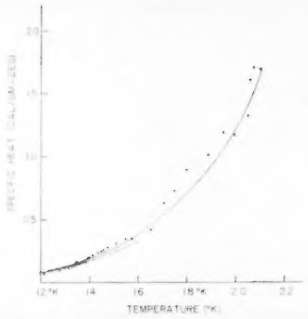


FIG. 10. Specific heat c_v *versus* temperature. Circles represent specific heat c_v of liquid helium II obtained using the thermal Rayleigh disk as a calorimeter. The solid line gives early data by Keesom *et al.* employing customary methods, while similarly the broken line combines data obtained more recently by Keesom and Westmijze and by Hull, Wilkinson, and Wilks.

In fact since v_2 is an accurately determinable single quantity, we have chosen in this instance to employ (23) conversely for computing values of specific heat c_v, in terms of our measured torques. Thus Fig. 10 gives specific heat c_v as a function of temperature T(°K) with our values, obtained using the thermal Rayleigh disk as a calorimeter, represented by circles ●. For comparison, specific heat data previously obtained by investigators using customary calorimetric methods are included. Thus the solid curve ———— was given by W. Keesom[25] on the basis of early measurements by Keesom and Keesom[26] and by Keesom and Clusius.[27] The broken curve on the other hand combines more recent data by Keesom and Westmijze[28] with that by Hull, Wilkinson, and Wilks.[29]

Our c_v results are probably least reliable in the central range of temperatures where accurate measurements of resonance sharpness Q are difficult; they are probably most reliable at lowest temperatures, where the relatively increased band widths and greater observed torques are conducive to accuracy.

(b) Low Temperature Dependence of Thermodynamic Combination $\rho_n/(ST)^2$

Returning to expression (7), these torque measurements provide a direct evaluation of the thermodynamic quantity $\rho_n/(ST)^2$ for the lower temperature regions, where $\rho_n/\rho_s \to \rho_n/\rho$. In Fig. 8(b) the log-log plot of this combination *versus* temperature shows a negative slope of approximately 8.5. This result deviates somewhat from Tisza's contention[1] that $\rho_n \sim S \sim T^{5.5}$ for the whole temperature range.

Landau[2] postulates a variable dependence on temperature approaching $\rho_n \sim T^4$, $S \sim T^3$ **as** $T \to 0$°K; thus the negative slope in Fig. 8(b) should presumably eventually approach the value 4.0. However we should not expect this to occur at these present temperatures. We do expect to observe the setting in of such trends during extension of this work on down to about 0.85°K, now in preparation with improved pumping facilities.

VII. CONCLUSIONS

The principal conclusions which may be drawn from this investigation are accordingly as follows:

1. The Rayleigh disk provides a direct method for detecting second sound by purely mechanical means.

2. The Rayleigh disk method provides "an absolute measure of second sound intensity dependent only upon such fundamentals as the kinetic energy density and the geometry."[4]

3. The observed results are in complete conformity with predictions based on the two-fluid concept. Although not necessarily providing a confirmation of that hypothesis, the results do substantiate an "ordered internal particle motion" of some sort associated with heat current in helium II.

4. On the basis of the two-fluid concept the results conform to theory throughout each of the following temperature intervals selected according to whether one fluid component, or the other, or both, provide the effective torque:

(a) Near the λ-point (2.1°K to the λ-point) superfluid should theoretically provide the dominant torque. Since the measured values agree well with those predicted for this range, the superfluid component appears to be contributing the predominant share. Accordingly the disk recognizes the motion of superfluid.

(b) At the lower temperatures (below about 1.75°K) normal fluid torque τ_n should theoretically predominate. Since agreement occurs for this range (down to 1.2°K) it may be concluded that the disk recognizes the motion of normal fluid.

(c) For intermediate temperatures (roughly 1.75°K to 2.1°K) both fluids should contribute significantly to the torque. Agreement for this range reveals the composite contributions by both fluids.

5. The appropriate individual and collective contributions by normal fluid and superfluid to the measured torque on the Rayleigh disk throughout the temperature range thus far investigated (1.2°K to the λ-point) appear to confirm the generalized Bernoulli relationship[4] [Eq. (20) of the present paper].

6. "In view of this present generalized form of the Bernoulli relationship, it would appear that for the majority of classical hydrodynamical devices (and experiments) there exists the thermal counterpart in liquid helium II."[9] The thermal pitot tube experiment constitutes another such application.

7. The thermal Rayleigh disk may be employed as a calorimeter, and measurements of c_v for liquid helium II thus determined are consistent with previously published values obtained by customary methods.

[25] W. Keesom, *Helium* (Elsevier Publishing Company, Inc., Houston, Texas, 1942), p. 215.
[26] W. Keesom and A. Keesom, Proc. Koninkl. Nederland, Akad. Wetenschap. **35**, 736 (1932); Physica **2**, 557 (1935).
[27] W. Keesom and K. Clusius, Proc. Koninkl. Nederland, Akad. Wetenschap. **35**, 307 (1932).
[28] W. Keesom and W. Westmijze, Physica **8**, 1004 (1941).
[29] Hull, Wilkinson, and Wilks, Proc. Phys. Soc. (London) **64**, 379 (1951).

Fermi Energy of Metallic Lithium

RICHARD A. SILVERMAN*
Department of Physics, Harvard University, Cambridge, Massachusetts
(Received October 8, 1951)

A boundary condition method is developed for deriving the coefficient E_{2n} in the power series expansion of the energy of an electron of wave number k moving in the lattice of an alkali metal. (The entire calculation proceeds within the framework of the Wigner-Seitz atomic sphere approximation.) If the electron wave function is expanded as $\psi_k(\mathbf{r})=e^{i\mathbf{k}\cdot\mathbf{r}}(u_0+u_1k+u_2k^2+\cdots)$ it is shown that the boundary condition $[(\partial/\partial r)(s \text{ part of } u_{2n})]_{r=r_s}=0$ leads naturally to an evaluation of E_{2n} in terms of values at r_s of homogeneous solutions of the Schrödinger equation and their derivatives with respect to energy and radius. In this way, a simple expression for E_4 is obtained analogous to that derived by Bardeen for E_2. For the case of metallic lithium, this expression leads to the value $E_4=-0.031$, which agrees with that obtained by the more tedious method of evaluating the expectation value of the Hamiltonian using a wave function correct to the second order in k.

INTRODUCTION

THE author and Kohn[1] have performed a Wigner-Seitz calculation of the coefficient E_4 in the power series expansion for the energy

$$E(k)=E_0+E_2k^2+E_4k^4+\cdots, \tag{1}$$

of an electron with wave number k, moving in the lattice of metallic lithium. The method consists in a straightforward evaluation[2] of the expectation value of the appropriate Hamiltonian using a wave function correct to the second order in k. The value of the cohesive energy obtained by a linear variation function method[1] was combined with several small corrections by Herring,[3] who obtained a theoretical value for the cohesive energy of lithium which agrees with the experimental value to within the probable error in either.

Although the linear variation function method gives more accurate results than those obtained from the power series (1), we present here a boundary condition method of obtaining the general coefficient E_{2n}, which avoids evaluating an expectation value of the Hamiltonian. In particular, we shall arrive quite directly at the expression for E_2 derived by Bardeen.[4]

BARDEEN'S PROCEDURE

If we write the wave function of an electron in the lattice of metallic lithium[5] as a power series in k

$$\psi_k(\mathbf{r})=e^{i\mathbf{k}\cdot\mathbf{r}}u_k(\mathbf{r})=e^{i\mathbf{k}\cdot\mathbf{r}}(u_0+u_1k+u_2k^2+\cdots), \tag{2}$$

(neglecting any dependence on the direction of $\mathbf{k}$, which is, of course, a fundamental weakness of this approach), the Schrödinger equation[6]

$$(-\nabla^2+V(r))\psi_k(\mathbf{r})=E(k)\psi_k(\mathbf{r}), \tag{3}$$

separates into the following set of equations familiar from ordinary perturbation theory:

$$\begin{aligned}
&(-\nabla^2+V-E_0)u_0=0,\\
&(-\nabla^2+V-E_0)u_1=2i(\partial u_0/\partial z),\\
&(-\nabla^2+V-E_0)u_2=2i(\partial u_1/\partial z)+(E_2-1)u_0,\\
&\quad\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&(-\nabla^2+V-E_0)u_{2n}=2i(\partial u_{2n-1}/\partial z)+(E_2-1)u_{2n-2}\\
&\qquad\qquad+\cdots+E_{2n}u_0,\\
&(-\nabla^2+V-E_0)u_{2n+1}=2i(\partial u_{2n}/\partial z)+(E_2-1)u_{2n-1}\\
&\qquad\qquad+\cdots+E_{2n}u_1,\\
&\quad\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned} \tag{4}$$

Here we have chosen $\mathbf{k}$ along the z axis of a spherical coordinate system. $V(r)$ is the empirical ion-core potential computed by Seitz.[7]

The procedure adopted by Bardeen[4] is to find particular solutions of the set (4) by inspection. These are then made to satisfy the Wigner-Seitz boundary condition[8] by adding appropriate multiples of solutions of the homogeneous equation

$$(-\nabla^2+V-E_0)\psi=0. \tag{5}$$

Consider, for example, the second equation of (4),

---

* Now at the Research Division, Philco Corporation, Philadelphia, Pennsylvania.

[1] R. A. Silverman and W. Kohn, Phys. Rev. **80**, 912 (1950).

[2] A numerical error was subsequently discovered which modifies the value of the entries of Table I of reference 1 to those given in an erratum, Phys. Rev. **82**, 283 (1951).

[3] C. Herring, Phys. Rev. **82**, 282 (1951).

[4] J. Bardeen, J. Chem. Phys. **6**, 367 (1938). Bardeen was concerned only with the construction of u_1, but his procedure is applicable to the construction of the general u_n, as we shall see below.

[5] The following considerations apply to any monovalent metal to the extent that there is available a radial ion-core potential which takes into account adequately the interaction between the closed-shell and valence electrons. This is certainly the case for lithium and sodium. See discussion in F. Seitz, *Modern Theory of Solids* (McGraw-Hill Book Company, Inc., New York, 1940), p. 348.

[6] Length is measured in units of the Bohr radius, energy in Rydberg units.

[7] F. Seitz, Phys. Rev. **47**, 400 (1935). Although Seitz used the correct ion-core potential in deriving the results reported in this reference, Herring discovered a small but significant error in the published potential. Bardeen kindly supplied us with the correct potential.

[8] It will be recalled that if f is even, the Wigner-Seitz boundary condition is $\partial f(r_s)/\partial r=0$, where r_s is the radius of the s-sphere, whereas if f is odd, the boundary condition is $f(r_s)=0$. The ground-state wave function u_0 is even, so that by the construction given below, u_i is even if i is even, odd if i is odd. The value E_0 occurring in (5) is determined by the boundary condition $\partial u_0(r_s)/\partial r=0$.

namely

$$(-\nabla^2+V-E_0)u_1=2i(\partial u_0/\partial z). \tag{6}$$

By inspection, $-izu_0$ is a particular solution of (6). The complete solution must satisfy the boundary condition $u_1(r_s)=0$; this can be achieved by adding an appropriate multiple of the p solution of (5). Thus, the desired solution of (6) is

$$u_1=iz(f_p/r-u_0), \tag{7}$$

where f_p is the radial part of the p solution of (5) which satisfies $f_p(r_s)=r_s u_0(r_s)$.

Now we place (7) in the expectation value[9]

$$\frac{\int_V (u_0+u_1k)^*(-\nabla^2+V-2ik\partial/\partial z+k^2)u_0 d\tau}{\int_V (u_0+u_1k)^* u_0 d\tau}, \tag{8}$$

which represents $E(k)$ correct to second order in k. (V = volume of the s-sphere.) By some ingenious manipulation, Bardeen was able to transform the resultant expression for E_2 into

$$E_2=\gamma(rf_p'/f_p)_{r=r_s}, \tag{9}$$

where $f_p'=\partial f_p/\partial r$[10] and $\gamma=(4\pi/3)r_s^3u_0^2(r_s)$; u_0 is normalized, i.e.,

$$\int_V u_0^2 d\tau=1.$$

We have discovered that the boundary condition

$$\left[\frac{\partial}{\partial r}(s \text{ part of } u_{2n})\right]_{r=r_s}$$

will lead automatically to the expression (9) for E_2, and to similar, more complicated expressions for E_4, E_6, etc. Let us first evaluate E_2 by this method.

EVALUATION OF E_2 BY BOUNDARY CONDITION METHOD

Consider the third equation of the set (4), namely

$$(-\nabla^2+V-E_0)u_2=2i\partial u_1/\partial z+(E_2-1)u_0. \tag{10}$$

If we operate on $-izu_1+\frac{1}{2}z^2u_0$ with $(-\nabla^2+V-E_0)$, we obtain $2i\partial u_1/\partial z-u_0$, which differs from the right-hand side of (10) only by the absence of the factor E_2u_0. Thus if f has the property that

$$(-\nabla^2+V-E_0)f=u_0, \tag{11}$$

a particular solution of (10) is $-izu_1+\frac{1}{2}z^2u_0+E_2f$. Such an f is given formally by differentiating $(-\nabla^2+V-E)u_0(E,r)=0$ with respect to E, and then setting $E=E_0$.[11] This gives

$$(-\nabla^2+V-E_0)(\partial u_0/\partial E)_{E=E_0}=u_0. \tag{12}$$

Thus,

$$-izu_1+\tfrac{1}{2}z^2u_0+E_2(\partial u_0/\partial E)_{E=E_0} \tag{13}$$

is a particular solution of (10). Recalling (7), it is seen that (13) is a linear combination of s and d functions. Therefore, the radial derivatives of the s and d parts of (13) must vanish at the surface of the s-sphere. To insure this we add appropriate multiples of s and d solutions of the homogeneous Eq. (4). Accordingly, the complete solution is

$$u_2=(\tfrac{2}{3}rf_p-\tfrac{1}{6}r^2u_0+c_df_d)P_2+(\tfrac{1}{3}rf_p-\tfrac{1}{6}r^2u_0 +E_2(\partial u_0/\partial E)_{E=E_0}+c_su_0)P_0. \tag{14}$$

Here P_0, P_2 are Legendre polynomials, and c_s, c_d are constants; f_d satisfies

$$\left(-\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)+V+\frac{6}{r^2}-E_0\right)f_d=0. \tag{15}$$

(Henceforth when $E=E_0$ and $r=r_s$ appear as subscripts, they will be abbreviated as E_0 and r_s, respectively.)

As far as the definition of $u_0(E,r)$ for values of E not equal to E_0 is concerned, we observe that $(\partial u_0(E,r)/\partial E)_{E_0}$ satisfies (10) provided the function $u_0(E,r)$ reduces continuously to $u_0(E_0,r)$ when E approaches E_0. Thus, if we define $u_0^n(E,r)$ as the solution of $(-\nabla^2+V-E)u_0^n(E,r)=0$ which is normalized for any E, $u_0(E,r)$ is given by $F(E)u_0^n(E,r)$, where $F(E)$ is any continuous function of E such that $F(E_0)=1$. Then we have

$$(\partial u_0(E,r)/\partial E)_{E_0}=F'(E_0)u_0^n(E_0,r) +(\partial u_0^n(E,r)/\partial E)_{E_0}, \tag{16}$$

i.e., any two definitions of $u_0(E,r)$ lead to values of $(\partial u_0(E,r)/\partial E)_{E_0}$ which differ only by a multiple of $u_0(E_0,r)$, a difference which is already within the arbitrariness of the functions u_n. [This fact is apparent from the form of (4).] Moreover, since $u_0'(E_0,r_s)=0$, it is clear that $(\partial^2u_0/\partial r\partial E)_{r_s,E_0}$ is uniquely defined.

Now we set the s part of (14) equal to zero:

$$\tfrac{1}{3}r_sf_p'(r_s)+E_2(\partial^2u_0/\partial r\partial E)_{r_s,E_0}=0, \tag{17}$$

i.e.,

$$E_2=-\tfrac{1}{3}r_sf_p'(r_s)/(\partial^2u_0/\partial r\partial E)_{r_s,E_0}. \tag{18}$$

The quantity $(\partial^2u_0/\partial r\partial E)_{r_s,E_0}$ can be evaluated by multiplying both sides of (12) by u_0, integrating over

[9] If $\psi^{(m)}$ and $\psi^{(n)}$ represent the true wave function to orders m and n, respectively, $\int\psi^{(m)*}H\psi^{(n)}d\tau/\int\psi^{(m)*}\psi^{(n)}d\tau$ represents the energy correct to order $m+n+1$.

[10] "Prime" will have this meaning hereafter.

[11] This procedure was originated by C. Herring and A. G. Hill in their paper on metallic beryllium, Phys. Rev. **58**, 132 (1940). These authors constructed an expression equivalent to Eq. (14).

the s-sphere, and invoking Green's theorem.[12] Thus,

$$\int_V u_0(-\nabla^2+V-E_0)(\partial u_0/\partial E)_{E_0}d\tau=\int_V u_0^2 d\tau=1,$$

$$\int_V (\partial u_0/\partial E)_{E_0}(-\nabla^2+V-E_0)u_0 d\tau + \int_s \{-u_0(\partial^2 u_0/\partial r\partial E)_{r_s,E_0} + (\partial u_0/\partial r)(\partial u_0/\partial E)_{E_0}\}d\sigma=1, \tag{19}$$

$$-4\pi r_s^2 u_0(r_s)(\partial^2 u_0/\partial r\partial E)_{r_s,E_0}=1,$$

$$(\partial^2 u_0/\partial r\partial E)_{r_s,E_0}=-1/4\pi r_s^2 u_0(r_s).$$

Substituting this value into (17), we find

$$E_2=(4\pi/3)r_s^3 u_0^2(r_s)(rf_p'/f_p)_{r_s}, \tag{20}$$

using $f_p(r_s)=r_s u_0(r_s)$. This is precisely the expression obtained by Bardeen as described above.

CONSTRUCTION OF u_n

We have seen that u_n must satisfy

$$\begin{aligned}(-\nabla^2+V-E_0)u_n&=2i\partial u_{n-1}/\partial z+(E_2-1)u_{n-2}\\&\quad+\cdots+E_n u_0 \quad (n \text{ even}),\\(-\nabla^2+V-E_0)u_n&=2i\partial u_{n-1}/\partial z+(E_2-1)u_{n-2}\\&\quad+\cdots+E_n u_1 \quad (n \text{ odd}).\end{aligned} \tag{21}$$

If we assume that u_n starts as

$$-\left[\frac{(iz)^n}{n!}u_0+\frac{(iz)^{n-1}}{(n-1)!}u_1+\cdots+iz u_{n-1}\right], \tag{22}$$

it can easily be shown that the difference between the right-hand side of (21) and the result of operating on (22) with $(-\nabla^2+V-E_0)$ consists only of terms involving solutions of the homogeneous Schrödinger equation (5) and their E derivatives of various orders evaluated at E_0. These terms can be generated by operating with $(-\nabla^2+V-E_0)$ on higher E derivatives of said homogeneous solutions.[13] Finally, suitable multiples of the homogeneous solutions themselves must be added to the sum of (22) and these E derivatives in order to satisfy the appropriate boundary condition at the surface of the s-sphere. For example, if we apply $(-\nabla^2+V-E_0)$ to (22) for the case $n=3$, we find that the difference between the resulting expression and the right-hand side of the third equation of (4) is $iE_2P_1f_p$; similarly, for $n=4$ the difference is $E_4u_0+E_2c_dP_2f_d+E_2^2(\partial u_0^n/\partial E)_{E_0}$. Accordingly, the complete solutions u_3 and u_4 are given by

$$\begin{aligned}u_3&=\tfrac{1}{6}iz^3u_0+\tfrac{1}{2}z^2u_1-izu_2+iE_2P_1(\partial f_p/\partial E)_{E_0}\\&\quad+c_pP_1f_p+c_fP_3f_f\\u_4&=-(1/24)z^4u_0+\tfrac{1}{6}iz^3u_1+\tfrac{1}{2}z^2u_0-izu_3\\&\quad+E_4(\partial u_0/\partial E)_{E_0}+\tfrac{1}{2}E_2^2(\partial^2 u_0^n/\partial E^2)_{E_0}\\&\quad+E_2c_dP_2(\partial f_d/\partial E)_{E_0}+\gamma_s u_0+\gamma_d P_2 f_d+\gamma_g P_4 f_g,\end{aligned} \tag{23}$$

where the c's and γ's are constants, P_l is the Legendre polynomial of degree l, and f_l is the radial solution of (6) with angular momentum quantum number l.

It will be observed that u_{2n} must contain the term E_{2n} in order to generate the $E_{2n}u_0$ term appearing in the right-hand side of (21). This accounts for the possibility of evaluating E_{2n} from the boundary condition

$$\left[\frac{\partial}{\partial r}(s \text{ part of } u_{2n})\right]_{r_s}=0. \tag{24}$$

For, since $u_0'(r_s)=0$, the equation obtained from (24) cannot be satisfied by appropriate choice of a multiple of u_0 added to u_{2n}. Inasmuch as $(\partial^2 u_0/\partial r\partial E)_{r_s,E_0}\neq 0$, the boundary condition (24) represents an equation which can be solved explicitly for E_{2n}.

EVALUATION OF E_4

We now give an outline of the algebra leading from the expression (23) for u_4 to a simple formula for E_4, via the boundary condition $[(\partial/\partial r)(s \text{ part of } u_4)]_{r_s}=0$. The algebra is straightforward but somewhat tedious; we will outline only the principal steps.

The radial derivative of the s part of u_4 evaluated at $r=r_s$ turns out to be

$$\begin{aligned}&\frac{1}{30}r_s^3u_0(r_s)-\frac{1}{10}r_s^2f_p(r_s)-\frac{1}{30}r_s^3f_p'(r_s)\\&\quad-\frac{1}{3}r_sE_2\left(\frac{\partial u_0}{\partial E}\right)_{r_s,E_0}-\frac{1}{6}r_s^2E_2\left(\frac{\partial^2 u_0}{\partial r\partial E}\right)_{r_s,E_0}\\&\quad-\frac{2}{15}r_sc_df_d(r_s)-\frac{1}{15}r_s^2c_df_d'(r_s)+\frac{1}{3}E_2\left(\frac{\partial f_p}{\partial E}\right)_{r_s,E_0}\\&\quad+\frac{1}{3}E_2r_s\left(\frac{\partial^2 f_p}{\partial r\partial E}\right)_{r_s,E_0}-\frac{i}{3}c_pf_p(r_s)-\frac{i}{3}r_sc_pf_p'(r_s)\\&\quad+\frac{1}{2}E_2^2\left(\frac{\partial^2 u_0^n}{\partial r\partial E^2}\right)_{r_s,E_0}+E_4\left(\frac{\partial^2 u_0}{\partial r\partial E}\right)_{r_s,E_0}.\end{aligned} \tag{25}$$

We are able to choose the following particularly convenient definition for $f_p(E,r_s)$ at points other than E_0:

$$f_p(E,r_s)=r_su_0^n(E,r_s). \tag{26}$$

It can easily be shown that the additional terms created by any other definition are canceled by the accompanying change in c_p. Thus we have

$$(\partial f_p/\partial E)_{r_s,E_0}=r_s(\partial u_0/\partial E)_{r_s,E_0}, \tag{27}$$

[12] This use of Green's theorem was suggested by W. Kohn.

[13] Note that in differentiating $(-\nabla^2+V-E)u_0^n=0$ more than once with respect to E, it is vital that the wave function be the normalized function u_0^n.

as well as $f_p(r_s)=r_s u_0(r_s)$. To reduce (25) further, we must evaluate c_p, which is determined by the boundary condition (p part of $u_3)_{r_s}=0$. The p part of u_3 can be found by reductions exactly analogous to those made in finding the s part of u_4. The result is

$$(p \text{ part of } u_3)=\frac{1}{10}ir^3u_0-\frac{3}{10}ir^2f_p-irE_2\left(\frac{\partial u_0}{\partial E}\right)_{E_0}$$
$$-\frac{2}{5}irc_df_d+iE_2\left(\frac{\partial f_p}{\partial E}\right)_{E_0}+c_pf_p, \quad (28)$$

which leads immediately to the following expression for c_p:

$$c_p=[ir_s/5f_p(r_s)][r_s^2u_0(r_s)+2c_df_d(r_s)]. \quad (29)$$

The value of c_d is obtained from

$$c_df_d'(r_s)=2E_2(\partial^2u_0/\partial r\partial E)_{r_s,E_0}, \quad (30)$$

which is an immediate consequence of comparing the radial derivatives of the s and d parts of u_2 [see Eq. (14)]. We have now reduced (25) to

$$\frac{1}{30}r_s^3u_0(r_s)\left(\frac{rf_p'}{f_p}\right)_{r_s}-\frac{1}{6}E_2r_s^2\left(\frac{\partial^2u_0}{\partial r\partial E}\right)_{r_s,E_0}$$
$$+\frac{1}{3}E_2r_s\left(\frac{\partial^2f_p}{\partial r\partial E}\right)_{r_s,E_0}+E_4\left(\frac{\partial^2u_0}{\partial r\partial E}\right)_{r_s,E_0}$$
$$+\frac{1}{2}E_2^2\left(\frac{\partial^3u_0^n}{\partial r\partial E^2}\right)_{r_s,E_0}+\frac{2}{15}E_2r_s^2\left(\frac{\partial^2u_0}{\partial r\partial E}\right)_{r_s,E_0}$$
$$\times\left[2\left(\frac{rf_p'}{f_p}\right)_{r_s}\left(\frac{rf_d'}{f_d}\right)_{r_s}^{-1}-1\right]. \quad (31)$$

The quantity $(\partial^3u_0/\partial r\partial E^2)_{r_s,E_0}$ can be evaluated in a fashion similar to that used to evaluate $(\partial^2u_0/\partial r\partial E)_{r_s,E_0}$. Multiply both sides of

$$\tfrac{1}{2}(-\nabla^2+V-E_0)(\partial^2u_0^n/\partial E^2)_{E_0}=(\partial u_0^n/\partial E)_{E_0}, \quad (32)$$

by u_0^n and integrate over the volume of the s-sphere, using Green's theorem. Then

$$\frac{1}{2}\int_V u_0^n(-\nabla^2+V-E_0)\left(\frac{\partial^2u_0^n}{\partial E^2}\right)_{E_0}d\tau$$
$$=\int_V u_0^n\left(\frac{\partial u_0^n}{\partial E}\right)_{E_0}d\tau=\frac{1}{2}\frac{d}{dE}\int_V u_0^{n2}d\tau=0, \quad (33)$$

or finally

$$(\partial^3u_0^n/\partial r\partial E^2)_{r_s,E_0}=0. \quad (34)$$

The evaluation of $(\partial^2f_p/\partial r\partial E)_{r_s,E_0}$ requires a little more care, but is otherwise quite straightforward. The analysis is as follows: if Φ_p denotes the numerical solution of the equation

$$\left(-\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)+V+\frac{2}{r^2}-E_0\right)\Phi_p=0. \quad (35)$$

Equation (26) requires that

$$f_p(r,E)=\frac{r_su_0(r_s,E)}{\Phi_p(r_s,E)}\Phi_p(r,E). \quad (36)$$

Introducing $P=r\Phi_p$, we obtain after some manipulation

$$\left(\frac{\partial^2f_p}{\partial r\partial E}\right)_{r_s,E_0}=\gamma^{-1}E_2\left(\frac{\partial u_0^n}{\partial E}\right)_{r_s,E_0}$$
$$+r_su_0(r_s,E_0)\frac{d}{dE}\left(\frac{P'(r_s,E)}{P(r_s,E)}\right)_{E_0}. \quad (37)$$

We have cast $(\partial^2f_p/\partial r\partial E)_{r_s,E_0}$ into this form because it is especially easy to evaluate numerically in terms of functions we have already tabulated in connection with our previous work.[1] (Note that $\{(d/dE)[P'(r_s,E)/P(r_s,E)]\}_{E_0}$ is easily converted into the simpler form

$$-\int_0^{r_s}P^2(E_0,r)dr/P^2(E_0,r_s).)$$

Setting (31) equal to zero, we finally find

$$\frac{E_2}{40\pi u_0(r_s)}+\frac{E_2}{24\pi u_0(r_s)}+\frac{1}{3}E_2r_s(\partial^2f_p/\partial r\partial E)_{r_s,E_0}$$
$$-\frac{E_4}{4\pi r_s^2u_0(r_s)}-\frac{E_2^2}{15\pi\gamma u_0(r_s)}(rf_d'/f_d)_{r_s}^{-1}$$
$$+\frac{E_2}{30\pi u_0(r_s)}=0, \quad (38)$$

or, combining terms and solving for E_4:

$$E_4=(2/5)r_s^2E_2-(4/15)E_2^2\gamma^{-1}r_s^2(rf_d'/f_d)_{r_s}^{-1}$$
$$+\gamma E_2(\partial^2f_p/\partial r\partial E)_{r_s,E_0}/u_0(r_s). \quad (39)$$

Substituting the appropriate numerical values into (39), we obtain $E_4=-0.031$, which represents a cohesive energy correction of $+0.5$ kcal/M. The discrepancy between the value of E_4 obtained by this method and that obtained by the expectation value method (-0.029) amounts to about 0.03 kcal/M, which is much less than experimental accuracy. Inasmuch as both methods yield E_4 as a difference of large numbers, with consequent loss of one significant figure, this discrepancy is easily accounted for despite the fact that our wave functions are known to be accurate to within less than one percent.

The author would like to express his indebtedness to Professors Walter Kohn and Harvey Brooks for their valuable advice in connection with this work.

Radiative Corrections to Compton Scattering

L. M. BROWN*
Cornell University, Ithaca, New York
AND
R. P. FEYNMAN†
California Institute of Technology, Pasadena, California

(Received August 6, 1951)

Corrections of order e^6 to the differential cross section for Compton scattering of unpolarized radiation by electrons are computed. The results for corrections ascribable to virtual photons are finite, relativistically invariant, and valid at all energies, but contain a term which depends logarithmically on an assumed small photon mass λ. A cross section of the same order has also been obtained for double Compton scattering in which one of the emitted photons has an energy small compared to the rest mass of the electron (with the electron initially at rest). This contains a term depending on $\ln\lambda$ which exactly compensates the similar term arising from virtual quanta in all observable cases. Approximations for low and high energies, as well as numerical results, are given. These disagree with results obtained previously by Schafroth.

THE object of this paper is to obtain the correction to the differential cross section for Compton scattering (Klein-Nishina formula) arising from the possibility that the electron may emit and reabsorb a virtual photon in connection with the scattering process. We shall apply the methods developed by one of us[1] to obtain an explicit cross section to order e^6 for unpolarized radiation, valid (in so far as the theory is valid) at all energies.

Previous workers have shown that the high frequency divergences which enter in the straightforward application of perturbation theory to this problem can be removed by charge and mass renormalization. Schafroth[2,3] has obtained a finite e^4-order matrix element in relativistic and gauge invariant form. He also showed, following the treatment of the analogous problem for scalar particles by Corinaldesi and Jost,[4] that the infrared divergence which occurs can be removed by addition of the double Compton cross section in which the incoming photon produces two photons on interacting with the electron, and he made explicit evaluation of the cross section (but not of the double scattering) in the nonrelativistic and extreme relativistic approximations. His results, however, disagree with ours in both limits.

Since the interpretation of any experiment to measure the radiative corrections requires a knowledge of the double Compton cross section, we have computed this also, for the case that one of the emitted photons has an energy in the laboratory system which is small compared to the electron rest energy.

After a brief introduction, we shall in Sec. II write down and discuss the matrix element for the corrections. Section III will detail the evaluation of the differential cross section. Section IV will be concerned with the infrared catastrophe and the double Compton effect. Sections V and VI will discuss limiting cases and some numerical results. Mathematical details will be reserved for the appendices.

The method of calculating this effect is given by Feynman,[5] and for brevity we will not repeat the discussion here but will simply carry out the explicit evaluation of the matrix elements involved. Our notation is that of reference 1.

Some improvement has been made in the method of computing matrix elements given in reference 1(b). This is described here in detail in Appendix Y.

I. THE KLEIN-NISHINA FORMULA FOR UNPOLARIZED RADIATION

The direct Compton effect, in which a photon of momentum q_1, polarization e_1, impinges on an electron of initial momentum p_1, to be scattered as a new photon of momentum q_2, polarization e_2, is represented by a matrix element

$$W=R+S \tag{1a}$$

with

$$R=\boldsymbol{e}_2(\boldsymbol{p}_1+\boldsymbol{q}_1-m)^{-1}\boldsymbol{e}_1, \quad S=\boldsymbol{e}_1(\boldsymbol{p}_1-\boldsymbol{q}_2-m)^{-1}\boldsymbol{e}_2. \tag{1b}$$

The final momentum of the electron is, of course, $\boldsymbol{p}_2=\boldsymbol{p}_1+\boldsymbol{q}_1-\boldsymbol{q}_2$. The terms correspond to the diagrams of Fig. 1.

We shall call

$$p_3=p_1+q_1=p_2+q_2, \quad p_4=p_1-q_2=p_2-q_1, \tag{2}$$

and define the important invariants κ, τ by

$$\begin{aligned} m^2\kappa &= m^2-p_3^2=-2p_1\cdot q_1=-2p_2\cdot q_2, \\ m^2\tau &= m^2-p_4^2=2p_1\cdot q_2=2p_2\cdot q_1. \end{aligned} \tag{3}$$

In the laboratory system, with ω_1 and ω_2 the energies of the incoming and outgoing photons, κ is $-2\omega_1/m$ and τ is $2\omega_2/m$. In terms of the quantities defined in

* Now at Northwestern University. Part of this work was done in partial fulfillment of the requirements for the Ph.D. degree at Cornell University.
† Now on leave at Centro Brasileiro de Pesquisas Fisicas, Rio de Janeiro, Brazil.
[1] R. P. Feynman, Phys. Rev. **76**, 749 (1949); and Phys. Rev. **76**, 769 (1949).
[2] M. R. Schafroth, Helv. Phys. Acta **22**, 501 (1949).
[3] M. R. Schafroth, Helv. Phys. Acta **23**, 542 (1950).
[4] E. Corinaldesi and R. Jost, Helv. Phys. Acta **21**, 183 (1948).
[5] This problem is discussed in reference 1, Appendix D, p. 788.

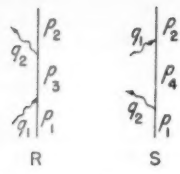


FIG. 1. Momentum diagrams for direct Compton effect.

(2) and (3) we have

$$m^2\kappa R=-\boldsymbol{e}_2(\boldsymbol{p}_3+m)\boldsymbol{e}_1, \quad m^2\tau S=-\boldsymbol{e}_1(\boldsymbol{p}_4+m)\boldsymbol{e}_2. \quad (4)$$

The differential cross section for the final photon to go into solid angle $d\Omega$, if the initial electron is at rest (laboratory system, $\boldsymbol{p}_1=m\boldsymbol{\gamma}_4$) is

$$d\sigma=e^4d\Omega(\omega_2^2/\omega_1^2)F \quad (5)$$

where F is the square of the matrix element of $W(1a)$,

$$F=|_2\langle W\rangle_1|^2. \quad (6)$$

If we are uninterested in the spin states of the electron, F may be replaced by $(2m^2)^{-1}U$ where

$$U=\tfrac{1}{4}\,\mathrm{Sp}[(\boldsymbol{p}_2+m)W(\boldsymbol{p}_1+m)\bar{W}]. \quad (7)$$

If, in addition, unpolarized radiation is used and the sum over polarization directions is required, $\boldsymbol{e}_1$ can be replaced by $\boldsymbol{\gamma}_\alpha$ and $\boldsymbol{e}_2$ by $\boldsymbol{\gamma}_\beta$ in the spur and half the sum over α, β taken (reference 1(b), Sec. 8). Then the term in (7) which is second order in R is

$$\tfrac{1}{2}(2m^2\kappa)^{-2}\,\mathrm{Sp}[(\boldsymbol{p}_2+m)\boldsymbol{\gamma}_\beta(\boldsymbol{p}_3+m)\boldsymbol{\gamma}_\alpha(\boldsymbol{p}_1+m)\boldsymbol{\gamma}_\alpha \times(\boldsymbol{p}_3+m)\boldsymbol{\gamma}_\beta]=4/\kappa^2-\tau/\kappa-2/\kappa. \quad (8)$$

The reduction can be accomplished by Eqs. (4a) and (36a) of reference 1(b). The term of second order in S is (8) with κ, τ interchanged, since S is obtained from R by replacing $\boldsymbol{p}_3$ by $\boldsymbol{p}_4$ after the average is taken on photon polarization. The cross term is

$$(2m^2\kappa)^{-1}(2m^2\tau)^{-1}\,\mathrm{Sp}[(\boldsymbol{p}_2+m)\boldsymbol{\gamma}_\beta(\boldsymbol{p}_3+m)\boldsymbol{\gamma}_\alpha \times(\boldsymbol{p}_1+m)\boldsymbol{\gamma}_\beta(\boldsymbol{p}_4+m)\boldsymbol{\gamma}_\alpha]=8/\kappa\tau-2/\tau-2/\kappa. \quad (9)$$

The sum gives for $U=m^2\sum_{\text{spin}}\sum_{\text{pol}}|_2\langle W\rangle_1|^2$:

$$U=4(\kappa^{-1}+\tau^{-1})^2-4(\kappa^{-1}+\tau^{-1})-(\kappa/\tau+\tau/\kappa) \quad (10)$$

and for the Klein-Nishina formula in terms of κ, τ we have:

$$d\sigma=\frac{2\pi e_4}{m^2}\left(\frac{\tau^2}{\kappa^2}\right)\left(\frac{d\tau}{\tau^2}+\frac{d\kappa}{\kappa^2}\right)U. \quad (11)$$

In the laboratory system, in view of $\kappa=-2\omega_1/m$, $\tau=2\omega_2/m$ and the Compton relation

$$\omega_1\omega_2(1-\cos\varphi)=m(\omega_1-\omega_2), \quad (12)$$

(11) can be written in the usual way

$$d\sigma=(e^4/2m^2)d\Omega(\omega_2^2/\omega_1^2)(\omega_1/\omega_2+\omega_2/\omega_1-\sin^2\varphi). \quad (13)$$

II. THE e^4-ORDER MATRIX ELEMENT

The diagrams of the first radiative corrections to term R of the Compton effect are given in Fig. 2. (See reference 1(b), Fig. 9.) The terms containing the analogous modifications of S can be obtained through-out by the interchange of $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, of $\boldsymbol{q}_1$ and $-\boldsymbol{q}_2$, and of $\boldsymbol{p}_3$ and $\boldsymbol{p}_4$. In the final result this means simply an interchange of κ and τ. Hence we need study only R, the S terms being obtained from the R terms immediately.

Terms N' and N'' give zero since there are no vacuum polarization effects for free photons.

Terms M' and M'' together give a factor[6] $r/2i$ times R, where

$$r=\ln(\Lambda/m)+9/4-2\ln(m/\lambda) \quad (14)$$

as shown in reference 1(b), Sec. 6. The quantity Λ is a temporary high frequency cutoff, introduced so that each diagram can be separately evaluated. The final result will become independent of Λ as $\Lambda\to\infty$. The "infrared catastrophe" discussed in Sec. IV is treated, at this point, by assuming the photons to have a small rest mass λ.

The term L is

$$L=\int \boldsymbol{e}_2(\boldsymbol{p}_3-m)^{-1}\boldsymbol{\gamma}_\mu(\boldsymbol{p}_3-\boldsymbol{k}-m)^{-1} \times\boldsymbol{\gamma}_\mu(\boldsymbol{p}_3-m)^{-1}\boldsymbol{e}_1k^{-2}d^4kC(k^2). \quad (15)$$

From this must be subtracted the mass correction for an electron travelling between the absorption and emission of the virtual quantum. Since (to order Δm)

$$(\boldsymbol{p}-m-\Delta m)^{-1}=(\boldsymbol{p}-m)^{-1}+(\boldsymbol{p}-m)^{-1}\Delta m(\boldsymbol{p}-m)^{-1},$$

this gives just the expression for L except that Δm replaces

$$\int \boldsymbol{\gamma}_\mu(\boldsymbol{p}_3-\boldsymbol{k}-m)^{-1}\boldsymbol{\gamma}_\mu k^{-2}d^4kC(k^2),$$

where Δm is the mass correction for cutoff Λ [reference 1(b), Eq. (21)]:

$$\Delta m=im[\tfrac{3}{8}+\tfrac{3}{2}\ln(\Lambda/m)]. \quad (16)$$

Since this diagram occurs for problems other than the one we consider here, we give the result in a general way. Each $(\boldsymbol{p}-m)^{-1}$ propagation factor has, as a consequence of diagrams like L, a correction to the first order in e^2 given by

$$(\boldsymbol{p}-m)^{-1}\int \boldsymbol{\gamma}_\mu(\boldsymbol{p}-\boldsymbol{k}-m)^{-1}\boldsymbol{\gamma}_\mu k^{-2}d^4kC(k^2)(\boldsymbol{p}-m)^{-1} -\Delta m(\boldsymbol{p}-m)^{-2}$$

$$=(4i)^{-1}\left\{(\boldsymbol{p}-m)^{-1}\left[\ln(\Lambda^2/m^2) +\frac{5}{2}-\frac{\eta}{\eta-1}+\frac{\eta(2-\eta)}{(\eta-1)^2}\ln\eta\right] -m(\boldsymbol{p}-m)^{-2}\left[\frac{\eta}{\eta-1}-\frac{\eta(3\eta-2)}{(\eta-1)^2}\ln\eta\right]\right\}, \quad (17)$$

where $m^2\eta=m^2-p^2$.

[6] The factor obtained in reference 1(b) is $-(e^2/2\pi)r$, but we have reserved a factor $e^2/\pi i$ for later inclusion.

Terms K' and K'' again possess a feature common to several problems, and we will therefore first discuss it in a general way. In all problems in which an electron interacts with a potential or a free or virtual photon there will be a piece of the diagram like Fig. 3. That is, there will be a partial factor in one of the matrix elements:

$$T=\int \gamma_\mu(\not p+\not q-\not k-m)^{-1}\boldsymbol{e}(\not p-\not k-m)^{-1} \times\gamma_\mu k^{-2}d^4kC(k^2). \quad (18)$$

It would be most convenient to have this evaluated in the general case of arbitrary $\boldsymbol{p}$ and $\boldsymbol{q}$. However, we have evaluated it only in the special case that $q^2=0$, $p^2=m^2$, with the matrix operating on a state u such that $\not p u=mu$. Calling $m^2\kappa=-2p\cdot q$ it is (Appendix Z):

$$\begin{aligned}8iT=4\kappa^{-1}[m^3\boldsymbol{e}+2\kappa^{-1}(e\cdot p)\not q]\int_{1-\kappa}^{1}\ln(1-v)dv/v\\ +2[(2m^2+\not p\not q-\not q\not p)\boldsymbol{e}\\ +2\kappa^{-1}(e\cdot p)(\not q+m\kappa)(3\kappa-2)(\kappa-1)^{-1}](\kappa-1)^{-1}\ln\kappa\\ +[2\ln(m^2/\Lambda^2)-1]m^2\boldsymbol{e}\\ -4(e\cdot p)[(\not q+m)(\kappa-1)^{-1}+\not q\kappa^{-1}]. \quad (19)\end{aligned}$$

If the final, rather than the initial, state is a free electron, the matrix required is $\bar T$, so the result is obtained directly from (19). For term K', this T for the case $\boldsymbol{p}=\boldsymbol{p}_1$, $\boldsymbol{q}=\boldsymbol{q}_1$, $\boldsymbol{e}=\boldsymbol{e}_1$ is to be multiplied on the left by $\boldsymbol{e}_2(\not p_1+\not q_1-m)^{-1}=\boldsymbol{e}_2(\not p_2-m)^{-1}$. Therefore, K' and the corresponding term K'' together give

$$K=K'+K''=(8i)^{-1}[\boldsymbol{e}_2(\not p_2-m)^{-1}T(p_1,q_1,e_1) +\bar T(p_2,q_2,e_2)(\not p_2-m)^{-1}\boldsymbol{e}_1]. \quad (20)$$

If we now examine the coefficients of the term $\ln(m^2/\Lambda^2)$ in K, L, and M, that is in (20), (17), and

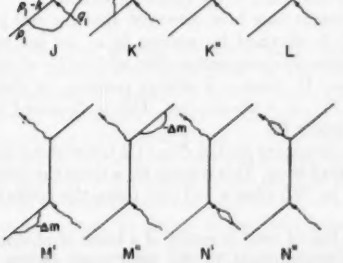


FIG. 2. Corrections to term R of Compton scattering.

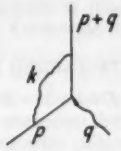


FIG. 3. Diagram for the expression T.

(14), we observe that K gives $(4m^2/8i)R$, L gives $(-2m^2/8i)R$, and M gives $(-2m^2/8i)R$. Therefore the terms dependent on Λ vanish. Since we shall find that the J integral is finite without cutoff, we note that the complete result is insensitive to Λ.

The term J is given by

$$J=\int \gamma_\mu(\not p_2-\not k-m)^{-1}\boldsymbol{e}_2(\not p_2-\not k-m)^{-1} \times\boldsymbol{e}_1(\not p_1-\not k-m)^{-1}\gamma_\mu k^{-2}d_4k. \quad (21)$$

For large k the factors in the integrand vary as k^{-n} with $n\geqq 5$ and the integration over k-space therefore converges. If we had included the convergence factor $C(k^2)$, the result would be independent of Λ as $\Lambda\to\infty$.

When the reciprocals are rationalized (e.g., $(\not p_2-\not k-m)^{-1}=(\not p_2-\not k+m)\cdot[(p_2-k)^2-m^2]^{-1}$), powers of k_μ up to the third appear in the numerator of the integrand. Therefore we shall have to evaluate integrals of the form:

$$\begin{aligned}J_{(0;\sigma;\sigma\tau;\sigma\tau\nu)}\\ =\int(1;k_\sigma;k_\sigma k_\tau;k_\sigma k_\tau k_\nu)[(p_2-k)^2-m^2]^{-1} \times[(p_2-k)^2-m^2]^{-1}[(p_1-k)^2-m^2]^{-1}k^{-2}d_4k. \quad (22)\end{aligned}$$

That is, for J_0 the factor $(1;k_\sigma;\cdots$ etc.) is replaced by unity, for J_σ by k_σ, for $J_{\sigma\tau}$ by $k_\sigma k_\tau$, and for $J_{\sigma\tau\nu}$ by $k_\sigma k_\tau k_\nu$. The manner in which J can be expressed in terms of these integrals is illustrated, for the case of matrix T in Appendix Z.

The J integrals can be worked out by the parametric methods described in reference 1(b) (Appendix). They involve integrals having four factors in the denominator and will lead, therefore, to integrals over three parameters. (J_0 is integrated in this manner in Appendix Y.) Generally these are very difficult to evaluate, although J_0 is particularly simple. This fact makes it possible to circumvent some of the difficulties of J_σ, $J_{\sigma\tau}$, and $J_{\sigma\tau\nu}$.

It is possible to express these other J integrals as linear combinations of the integral J_0 and of other integrals, all of which involve only three quadratic factors in the denominator. These latter, in parametric form, require only two parameters (and are much more easily evaluated than a direct attack on J_σ, say, would

indicate). This technique is useful in other problems also[7] and is described in detail in Appendix Y.

III. CROSS SECTION FOR UNPOLARIZED LIGHT

If we call the sum $J+K+L+M=R^{(1)}$, then $(e^2/\pi i)R^{(1)}$ will be the correction to the matrix R of the direct effect (1). If the corresponding correction to the term S is called $(e^2/\pi i)S^{(1)}$, the corrected matrix for the Compton effect is

$$W'=W+(e^2/\pi i)W^{(1)}=R+S+(e^2/\pi i)(R^{(1)}+S^{(1)}). \quad (23)$$

The absolute square of the matrix element of W', taken between the initial and final electron states, gives the probability of transition correct to one order in e^2 higher than (6). We shall calculate in this paper only the cross section averaged over spin directions of the electron and polarization directions of the photons.

We need the spur:

$$\tfrac{1}{4}\,\mathrm{Sp}[(\not p_2+m)W'(\not p_1+m)\bar W'] \quad (24)$$

as in (7). Considering terms up to the first order in e^2 (which are all that are valid), (24) is

$$U-\tfrac{1}{4}\{(e^2/\pi i)\,\mathrm{Sp}[(\not p_2+m)W(\not p_1+m)\bar W^{(1)}] -(e^2/\pi i)\,\mathrm{Sp}[(\not p_2+m)W^{(1)}(\not p_1+m)\bar W]\}. \quad (25)$$

In evaluating (25) for unpolarized light we have replaced $\boldsymbol{e}_1$ by $\boldsymbol{\gamma}_\alpha$ and $\boldsymbol{e}_2$ by $\boldsymbol{\gamma}_\beta$ and taken one-half of the resulting sum as discussed in connection with (8). Some algebraic details are discussed in Appendix Z.[8]

The last two spurs in (25) are complex conjugates, so that the correction to U is $-e^2/\pi$ times the real part of

$$U^{(1)}=-(4i)^{-1}\,\mathrm{Sp}[(\not p_2+m)W^{(1)}(\not p_1+m)\bar W]. \quad (26)$$

That is, U is to be replaced in (11) by

$$U'=U-(e^2/\pi)\mathbf{R.P.}U^{(1)}. \quad (27)$$

If we let

$$P(\kappa,\tau)=-(4i)^{-1}\,\mathrm{Sp}[(\not p_2+m)R^{(1)}(\not p_1+m)\bar W] \quad (28)$$

then

$$U^{(1)}=P(\kappa,\tau)+P(\tau,\kappa) \quad (29)$$

since the $S^{(1)}$ diagrams are obtained from the $R^{(1)}$ diagrams (for unpolarized light) by the interchange of $\not p_2$ and $\not p_1$ and of $\not q_1$ and $-\not q_2$; hence the final result, simply by interchange of κ and τ.

The final result obtained in this way is:

$$\begin{aligned}P(\kappa,\tau)=&(1-2y\,\mathrm{ctnh}2y)\,\ln\lambda\cdot U\\ &-2y\,\mathrm{ctnh}2y[2h(y)-h(2y)]U\\ &+[-4y\sinh 2y(\kappa\tau)^{-1}(2-\cosh 2y)\\ &+2y\,\mathrm{ctnh}y]h(y)+\ln\kappa\Big\{4y\,\mathrm{ctnh}2y\Big[\frac{4}{\kappa\tau}\cosh^2 y\\ &+\frac{\kappa-6}{2\tau}\,\mathrm{sech}2y+\frac{4}{\kappa^2}+\frac{1}{\kappa}+\frac{\tau}{2\kappa}+\frac{\kappa}{\tau}-1\Big]\\ &+\frac{3\tau}{2\kappa^2}+\frac{3\tau}{2\kappa}+\frac{3}{\tau}+1-\frac{7}{\kappa\tau}+\frac{8}{\kappa}+\frac{8}{\kappa^2}+\frac{2\kappa-\tau^2-\kappa^2\tau}{2\kappa^2\tau(\kappa-1)}\\ &-\frac{1}{2\tau}\frac{2\kappa^2+\tau}{(\kappa-1)^2}\Big\}+y^2\,\mathrm{csch}^2 y\Big[\frac{2}{\kappa}-\frac{7}{4}-\frac{\kappa}{4}-\frac{3}{4}\frac{\tau^2}{\kappa}\Big]\\ &-4y\tanh y\Big(\frac{1}{2}-\frac{1}{\kappa}\Big)+4\Big(\frac{1}{\kappa}+\frac{1}{\tau}\Big)^2\\ &-\frac{12}{\kappa}-\frac{3\kappa}{2\tau}-2\frac{\kappa}{\tau^2}+\frac{1}{\kappa-1}\Big(\frac{\kappa}{\tau}+\frac{1}{2}\Big)\\ &+G_0(\kappa)\Big[\frac{\kappa^2}{\tau}+\frac{\tau}{\kappa^2}+\frac{\kappa}{\tau}+\kappa+\frac{1}{2}-\tau+\frac{2}{\kappa}-\frac{3}{\tau}-1\Big]\\ &+\text{terms antisymmetric in }\kappa,\tau, \quad (30)\end{aligned}$$

where

$$4\sinh^2 y=-(\kappa+\tau) \quad (30a)$$

$$h(y)=y^{-1}\int_0^y u\,du\,\mathrm{ctnh}u \quad (30b)$$

$$G_0(\kappa)=-2\kappa^{-1}\int_{1-\kappa}^1 \ln(1-u)\,du/u. \quad (30c)$$

This is to be added to the same expression with κ and τ interchanged (29) and the real part taken to get the correction to the Klein-Nishina formula (11). We discuss this result in the following sections.

We might note here, however, that the real part of $P(\kappa,\tau)$ is obtained by writing $\ln|\kappa|$ for $\ln\kappa$ and by writing for $G_0(\kappa)$ expression (30c) with $\ln(1-u)$ replaced by $\ln(u-1)$. Since τ is always positive, on the other hand, $P(\tau,\kappa)$ is always real. This is discussed further in Appendix W.

The imaginary part of $P(\kappa,\tau)$ is not without interest, as we shall show. This is given by π times the coefficient of $\ln\kappa$ in (30) plus $\pi\ln(1-\kappa)$ times the coefficient of $G_0(\kappa)$.

The loss of total intensity of a beam of photons is of course proportional to the total cross section for a photon to be scattered out of the beam. But this

[7] It has been applied by G. R. Lomanitz to completely evaluate the e^6 corrections to the Möller scattering cross section of electrons in his thesis *Second Order Effects in the Electron-Electron Interaction*, Cornell, 1950. Again, in the problem of scattering of light by light, the integral with unit numerator is easily done, and the other integrals can be reduced to it and simpler integrals algebraically. But here the algebraic complexity makes the problem extremely tedious.

[8] In actual evaluation it was found easier to take the spur first and perform the integrals later. Thus, in place of the expression T (Eq. 18), the expression $\boldsymbol{T}$ (Eq. A41) was substituted and the values of the integrals from Appendix X substituted after taking the spur. This has the advantage that some of the integrals do not appear, or appear only in simpler combinations.

decrease in forward intensity is the result of an interference between the incident photon and a photon scattered exactly in the forward direction. Therefore, as is well known, the imaginary part of the forward scattering amplitude is proportional to the total cross section (formally this is referred to as the unitary property of the S-matrix). We can use this relation to check the imaginary part of $P(\kappa, \tau)$ for the case of zero scattering angle (for which, of course, $p_1=p_2$, $q_1=q_2$, $\kappa=-\tau$).

We write $P(\kappa, \tau)$ again as a sum

$$P(\kappa, \tau)=(4m^2i)\tfrac{1}{4}\sum_{\text{spin}}\sum_{\text{pol}}{}_1\langle R^{(1)}\rangle_2\,{}_2\langle W\rangle_1, \quad (31)$$

and can show easily that ${}_2\langle W\rangle_1=i/m$ if there is no spin change and no polarization change, and zero otherwise. This can be seen, aside from the phase factor i, from the fact that for small scattering angles the Klein-Nishina formula (13) is $d\sigma=r_0^2d\Omega$ in the laboratory system. Since ${}_1\langle R^{(1)}\rangle_2$ also vanishes when ${}_2\langle W\rangle_1$ does,

$$P(\kappa, \tau)=m\sum_{\text{spin}}\sum_{\text{pol}}{}_1\langle R^{(1)}\rangle_2. \quad (32)$$

But, including all factors, the complete e^4-order matrix element of $R^{(1)}$ is (according to reference 1(b)):

$$X=\frac{e^2}{\pi i}\cdot\frac{2\pi e^2}{(m\tau/2)}\cdot\frac{1}{4}\sum_{\text{spin}}\sum_{\text{pol}}{}_1\langle R^{(1)}+S^{(1)}\rangle_2 \quad (33)$$

and from the unitary property referred to above, it follows that the *total* cross section for Compton scattering to order e^4 is just twice the real part of X. Therefore,

$$\sigma_{\text{total}}(\text{to order } e^4)=2\mathbf{R}.\mathbf{P}.X=(2r_0^2/\tau)\mathbf{I}.\mathbf{P}.P(\kappa, \tau) \quad (34)$$

since $S^{(1)}$ has no imaginary part. That (30) satisfies this identity can be readily verified.[9]

IV. THE INFRARED CASTASTROPHE AND THE DOUBLE COMPTON EFFECT

In Sec. II we have derived the differential cross section for Compton scattering for unpolarized light, including radiative corrections, to order e^6. The cross section took the form

$$d\sigma=d\sigma_{\text{K.N.}}[1+(e^2/\pi)\delta] \quad (35)$$

with

$$\delta=-U^{(1)}/U. \quad (35a)$$

There are two reasons why this result cannot be compared directly with experiment. In the first place $U^{(1)}$ depends on the quantity λ to which no experimental significance has been attached. In the second place, it is impossible in principle to design an experiment which will guarantee that one and only one photon is emitted by the electron in the scattering process. The best one can do in an experiment is to require that if a second photon is emitted, its energy is less than some value $k_{\max}$. This can be done, for example, by measuring the energies of the final electron and photon to some specified accuracy, the sum of the errors in the measurement being less than $k_{\max}$. In such an experiment one would be measuring the cross section (35) plus the cross section for the double Compton effect, $d\sigma_D$, integrated over all possible directions of the second quantum and over its energy up to $k_{\max}$.

These two difficulties, both related to quanta of low energy (if $k_{\max}$ is small), in one case virtual, in the other real, are actually related. That this should be so, can be seen physically from the fact that it is difficult to distinguish between virtual and real quanta of extremely low energy since, by the uncertainty principle, a measurement made during a finite time interval will introduce an uncertainty in the energy of the quantum, which may enable a virtual quantum to be detected as a real one. It turns out in fact that $d\sigma_D$, integrated to $k_{\max}$, also contains an infrared divergence which just cancels the similar divergence in the radiative corrections. We are computing, of course, only to order e^6, but the multiple Compton scattering of a given higher order will also cancel all the radiative infrared catastrophes of the same order.

The problem is analogous to the perturbation theory treatment of the scattering of an electron by a potential, which has been considered by many workers, except that in our case the primary process is the Compton scattering considered in Sec. I. The cross section for emission of an additional photon $\boldsymbol{q}$ of energy ω goes for small ω as $(d^3q/\omega)(p_2/p_2\cdot q-p_1/p_1\cdot q)^2$ times the Klein-Nishina formula. Since this diverges as ω approaches zero, the probability of a single Compton process unaccompanied by such emission is zero. What is experimentally measured, however, is the probability that a Compton process occurs and that no other free photon is emitted except for a class of photons inaccessible to the experiment. This is equal, to our order of calculation, to the probability of the single process plus the probability of a double process in which one of the photons emitted is in the inaccessible class. This class is, of course, determined by the design of the experiment and a single calculation cannot suffice for all experiments. However, one feature common to all experiments will be a finite energy resolution, so that a part of the excluded class must consist of photons whose energy is less than some energy $k_{\max}$.

We will first therefore find that part of the differential cross section for double Compton scattering which gives rise to an infrared divergence. This will be integrated over all the directions of one of the photons, and over its energy from zero to a value $k_{\max}$, which we shall assume is small compared to the electron mass, and added to the previously obtained corrected cross section for single Compton scattering. It has already been pointed out that Schafroth[2] has demonstrated that a cancellation of the infrared divergence occurs in order e^6 when the double Compton cross section is added to

[9] W. Heitler, *The Quantum Theory of Radiation* (Oxford University Press, London, 1944), p. 157, Eq. (53). Our Eq. (33) agrees with this result with τ replacing Heitler's 2γ.

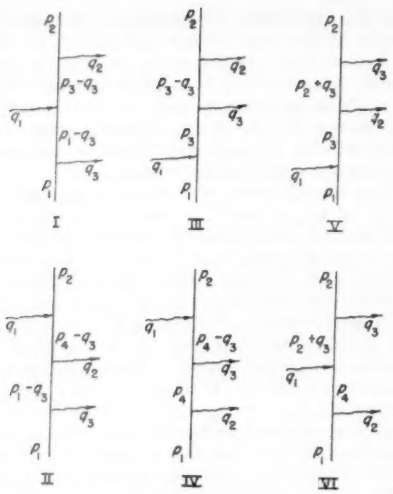


FIG. 4. Diagrams for the double Compton effect. Here $p_3=p_1+q_1$, $p_4=p_1-q_2$; the momentum condition is $p_1+q_1=p_2+q_2+q_3$.

the single scattering cross section, but we must obtain at least the zero order term in $k_{\max}$ (for $k_{\max}\ll m$) to obtain a useful result. The completely differential cross section for the double scattering has been computed by Eliezer,[10] but since we wish to make approximations and carry out an integration, it is simpler for us to obtain the desired cross section from the beginning.

Figure 4 gives the diagrams necessary for computing the cross section for double Compton scattering. The photon momentum $q_3=(\omega_3, \mathbf{q}_3)$ is assumed to be small in the following ($\omega_3\ll m$). This, of course, implies a definite coordinate system. To obtain a finite result we assume the photon has a small rest mass λ, so that $q^2=\lambda^2$. Keeping terms only to order ω_3^{-1}, we neglect q_3 occurring in the numerator of the (rationalized) matrix element terms, and terms of order ω_3 compared to $p_3^2-m^2$ and $p_4^2-m^2$ in the denominators.

We find that terms III and IV are not of the desired order. In view of the fact that we are to make matrix elements between the free electron states u_1, and u_2, a factor $(\not{p}_1+m)\not{e}_3$ (with e_3 the polarization vector of q_3) operating on the left of u_1, is equivalent to $2p_1\cdot e_3$ and a factor $\not{e}_3(\not{p}_2+m)$ operating on the right of u_2 is equivalent to $2p_2\cdot e_3$. Thus, with q_3 small, we get

$$\begin{aligned}\langle \mathrm{I}\rangle &= -Rp_1\cdot e_3/p_1\cdot q_3, \quad \langle \mathrm{II}\rangle = -Sp_1\cdot e_3/p_1\cdot q_3,\\ \langle \mathrm{V}\rangle &= Rp_2\cdot e_3/p_2\cdot q_3, \quad \langle \mathrm{VI}\rangle = Sp_2\cdot e_3/p_2\cdot q_3.\end{aligned} \tag{36}$$

Adding these we find the matrix for the double Compton process:

$$(R+S)\left(\frac{p_2\cdot e_3}{p_2\cdot q_3}-\frac{p_1\cdot e_3}{p_1\cdot q_3}\right), \quad |\mathbf{q}_3|\ll m. \tag{37}$$

Taking the absolute square of (37) and averaging over polarizations and spins in the usual manner, it is clear that we obtain the Klein-Nishina cross section $d\sigma_{\mathrm{K.N.}}$ (Eq. (11)) multiplied by the following factors: (a) $d^3\mathbf{q}_3/(2\pi)^3$, the density of states for $\boldsymbol{q}_3$ (neglecting its effect on the momentum balance, and therefore assuming it is emitted independently of $\boldsymbol{q}_2$), (b) e^2, from the additional interaction vertex, (c) $2\pi/\omega_3$, the normalization factor for the photon $\boldsymbol{q}_3$, (d)

$$\sum_{\mathrm{pol}}\left(\frac{p_2\cdot e_3}{p_2\cdot q_3}-\frac{p_1\cdot e_3}{p_1\cdot q_3}\right)^2=-\left(\frac{p_2}{p_2\cdot q_3}-\frac{p_1}{p_1\cdot q_3}\right)^2.$$

We collect these factors and integrate the photon momentum over all angles and from $\mathbf{q}_3=0$ to the sphere $|\mathbf{q}_3|=k_{\max}$, where $k_{\max}\ll m$. Thus,[11]

$$d\sigma_D=-\frac{e^2}{(2\pi)^2}d\sigma_{\mathrm{K.N.}} \times\int_{|\mathbf{q}_3|=0}^{|\mathbf{q}_3|=k_{\max}}\left(\frac{p_2}{p_2\cdot q_3}-\frac{p_1}{p_1\cdot q_3}\right)^2\frac{d^3\mathbf{q}_3}{(\mathbf{q}_3^2+\lambda^2)^{\frac{1}{2}}}. \tag{38}$$

Observe that if we replace $d\sigma_{\mathrm{K.N.}}$ by the cross section $d\sigma_0$ for an arbitrary process of one electron, the result (37) is valid for that process with one additional photon of small energy in the final state. For if the "small" photon $\boldsymbol{q}_3$ is emitted from an electron line of momentum $\boldsymbol{p}$ where $\boldsymbol{p}^2\neq m^2$, its effect will be negligible. The only diagrams contributing to the process with $\boldsymbol{q}$ emitted will be those of the original process, modified by the emission of $\boldsymbol{q}$ either before or after the original process. Thus the factor of $(R+S)$ in (37) will always be the factor modifying the original matrix element (for small $\boldsymbol{q}$) and the result (38) is the general result for an arbitrary process of one electron, $d\sigma_0$ replacing $d\sigma_{\mathrm{K.N.}}$.

It is most convenient to impose our restriction $k_{\max}\ll m$ in the laboratory system.[12] In this case, the result of integrating (38) (expressed in terms of invariants) is

$$d\sigma_D=-(e^2/\pi)d\sigma_{\mathrm{K.N.}}\{2(1-2y\,\mathrm{ctnh}2y)[\ln(2k_{\max}/\lambda) -\tfrac{1}{2}]+4y\,\mathrm{ctnh}2y[h(2y)-1]\}, \tag{39}$$

with y and $h(y)$ defined in (30).

[10] C. J. Eliezer, Proc. Roy. Soc. (London) **A187**, 210 (1946). When the class of photons inaccessible to the experiment does not consist simply of those below a given very small energy $k_{\max}$ (but consists, for example, of those in a given solid angle, or with a limited momentum component, or having energies too large to permit the approximations we have made) the contribution which these events make to the measured cross section can be obtained from Eliezer's formula. Explicitly, one must add to our result (39) the cross section for the double process given by Eliezer, integrated over all the photons in the class inaccessible to the experiment but which also exceed some arbitrary very small energy $k_{\max}$. The sum, of course, will not depend on $k_{\max}$.

[11] This expression (with $\lambda=0$) has been obtained previously. See, for example, R. Jost, Phys. Rev. **72**, 815 (1947), and F. Bloch and A. Nordsieck, Phys. Rev. **52**, 54 (1937).

[12] J. Schwinger, Phys. Rev. **76**, 790 (1949) has integrated (38) for the case of the scattering of an electron in a potential.

When (39) is added to (35) the effect is to replace the quantity

$$\{2(1-2y\,\mathrm{ctnh}2y)\ln\lambda-4y\,\mathrm{ctnh}2y[2h(y)-h(2y)]\}U$$

in $U^{(1)}$ by

$$\{2(1-2y\,\mathrm{ctnh}2y)[\ln(2k_{\max})-\tfrac{1}{2}]+8y\,\mathrm{ctnh}2y[h(2y)-h(y)-\tfrac{1}{2}]\}U. \quad (40)$$

We have now arrived at a physically understandable result as the quantity $k_{\max}$ which replaces λ in $U^{(1)}$ is the sum of the experimental uncertainties in the measurement of the final energies of the Compton scattered photon and electron. Our result can be compared with an experiment providing the energy resolution[10] is known and $k_{\max}$ is sufficiently small.

Under the limits of validity of our formula, the term in (40) containing $k_{\max}$ is positive, and thus makes a negative contribution to the cross section. As the energy resolution of an experiment improves, it is thus found that the measured Compton cross section gets smaller. This is reasonable since we are eliminating from our observations more double Compton events.

The expression (35a) for δ with $U^{(1)}$ given by (29) and U given by (10) is valid also for the correction to the two-quantum pair annihilation and the two-quantum pair production processes, provided that for the former problem we replace κ by $-\kappa$ and for the latter τ by $-\tau$. This occurs because in writing down the matrix element we represent the emission of a photon by $-q$ and its absorption by $+q$, and because a matrix $\not{p}$ representing an electron also represents a positron of four-momentum $-p$. However, the infrared divergences in these problems are not compensated by the corresponding three-quantum processes[11] (which are not divergent) but by the effect of Coulomb interaction. This will not be discussed further in this paper.

V. EXTREME RELATIVISTIC LIMIT

The Compton formula (12) can be written in the laboratory system, with scattering angle φ, as

$$(\kappa+\tau)/\tau=\tfrac{1}{2}\kappa(1-\cos\varphi). \quad (41)$$

We assume $|\kappa|\gg 1$ and consider the three cases listed in Table I. This table also lists the approximations made in obtaining the formulas for U and $U^{(1)}$ for the three cases and the corresponding conditions on the laboratory and center-of-mass scattering angles. The energies (in units of mc^2) of the incoming and outgoing quanta are represented in the laboratory system by ω_1 and ω_2, respectively, and in the c.m. system by ν. The results are as follows:

Case I.

$$U=2$$

$$U^{(1)}=4(1-2y\,\mathrm{ctnh}2y)\ln\lambda-8y\,\mathrm{ctnh}2y[2h(y)-h(2y)]+4yh(y)\,\mathrm{ctnh}y+\ln|\kappa|(4y\tanh y-1)-2y^2-4y\tanh y+3-(\ln|\kappa|)^2-\pi^2/6. \quad (42)$$

[11] Note that the two quantum pair processes are symmetric with respect to interchange of p_1 and p_2 so that (37) vanishes.

Case II.

$$U=-[(\kappa/\tau)+(\tau/\kappa)]$$

$$U^{(1)}=U\{(1-2y)(\tfrac{3}{2}+2\ln\lambda)+2y^2-\pi^2/6\}+\left(1+\frac{\tau}{2\kappa}+\frac{\kappa}{\tau}\right)\left\{\left[\ln\left(1+\frac{\tau}{\kappa}\right)\right]^2-\ln\left(1+\frac{\tau}{\kappa}\right)+2\ln\frac{\tau}{|\kappa|}\right\}+\left(1+\frac{\tau}{\kappa}+\frac{\kappa}{2\tau}\right)\left\{\left(\ln\left|1+\frac{\kappa}{\tau}\right|\right)^2-\ln\left(1+\frac{\tau}{\kappa}\right)-\ln\left|\frac{\tau}{\kappa}\right|+\pi^2\right\}. \quad (43)$$

Case III.

$$U=-\kappa/\tau$$

$$U^{(1)}=U\left\{2(1-2y)\ln\lambda+\ln\tau\left[2y-\frac{3}{2}\frac{\tau+1}{\tau}+\frac{1}{2\tau(\tau-1)}\right]-\frac{\pi^2}{3}-G_0(\tau)\left(\frac{1}{\tau}+\frac{\tau}{2}\right)+\frac{3}{2}+\frac{2}{\tau}\right\}. \quad (44)$$

The corrected cross sections, in the relativistic limit are given in the laboratory system by

$$d\sigma=(r_0^2/2)d\Omega(\tau^2/\kappa^2)[U-(e^2/\pi)U^{(1)}] \quad (45)$$

and in the c.m. system by

$$d\sigma=(r_0^2/8\nu^2)d\Omega[U-(e^2/\pi)U^{(1)}] \quad (46)$$

with $r_0=e^2/mc^2$.

As we have explained in the previous section, for actual comparison with experiment one must add to (45) the cross section for double Compton scattering (35) which is valid, of course, only in the laboratory system with $k_{\max}\ll m$. If we write (39) as

$$d\sigma_D=(-e^2/\pi)(r_0^2/2)d\Omega(\tau^2/\kappa^2)U_D \quad (47)$$

then we must replace $U^{(1)}$ in (45) and (46) by $U^{(1)}+U_D$. For our three cases, we get:

TABLE I. Approximation made in obtaining the extreme relativistic limit of (30), expressed in invariants, laboratory system quantities, and c.m. system quantities.

Case	Defined by	Leads to	Lab system conditions	c.m. conditions
I	$\lvert(\kappa+\tau)/\tau\rvert\ll 1$	$\lvert\kappa\rvert\approx\tau\gg 1$	$\omega_2\approx\omega_1$, $1-\cos\varphi\ll 1/\omega_1$ (φ near 0)	$\tan^2(\theta/2)\ll 1$
II	$\lvert(\kappa+\tau)/\tau\rvert\sim 1$	$\left.\begin{matrix}\tau\gg 1\\ \lvert\kappa+\tau\rvert\gg 1\end{matrix}\right\}$	ω_2 near $\omega_1/2$, $1-\cos\varphi\sim 1/\omega_1$ (φ near $(2/\omega_1)^{\frac{1}{2}}$)	$\nu^2\cos^2(\theta/2)\gg 1$, $\nu^2\sin^2(\theta/2)\gg 1$
III	$\lvert(\kappa+\tau)/\tau\rvert\gg 1$	$\left.\begin{matrix}\lvert\kappa+\tau\rvert\gg 1\\ \lvert\kappa\rvert\gg\tau\end{matrix}\right\}$	ω_2 near 1, $1-\cos\varphi\gg 1/\omega_1$ ($\varphi\gg(2/\omega_1)^{\frac{1}{2}}$)	$\tan^2(\theta/2)\gg 1$

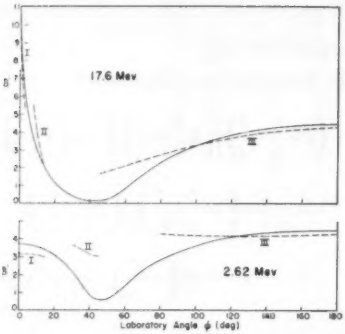


FIG. 5. Plot of $\delta' = -U^{(1)}/U - 2(1-2y\,\mathrm{ctnh}2y)\ln\lambda$ for 2.62 Mev and 17.6 Mev as calculated from the exact expressions (29) and (30) (solid curve). The dotted curves are calculated from the extreme relativistic formulas (42), (43), and (44), and are numbered accordingly.

Case I.

Same as (39).

Cases II and III.

$$U_D = U\{2(1-2y)[\ln(2k_{\max}/\lambda) - \tfrac{1}{2}] + 4y[y + (\pi^2/24y) - 1]\}. \quad (48)$$

In Fig. 5 we have plotted a comparison of the exact expression for $-U^{(1)}/U$ (leaving out the term proportional to $\ln\lambda$) with the limiting cases expressed by (42), (43), and (44). That is, if we write $U^{(1)} = a\ln\lambda + b$, we have plotted $-b/U$. Especially simple formulas result for b in the extreme relativistic limit for the cases $\varphi = 0$ and $\varphi = 180°$. At zero angle (where incidentally, the double Compton effect and the $\ln\lambda$ term in $U^{(1)}$ vanish):

$$b(0°) = -\ln\tau(1+\ln\tau) + 1.355. \quad (49)$$

At 180 degrees,

$$b(180°) = -4.225U. \quad (50)$$

Another simple case results from the condition $\kappa = -2\tau$, in the extreme relativistic limit. This corresponds to 90° scattering in the c.m. system. Here we get $U = 5/2$ and

$$b(90°, \text{c.m.}) = (2y^2 - 3y - 2.29)U. \quad (51)$$

TABLE II. Percent correction to the Compton cross section for unpolarized light arising from Eq. (30), excluding the term proportional to $\ln\lambda$, at zero degrees and at ninety degrees in the c.m. system; computed as a function of the laboratory energy of the incident photon from the special equations (49) and (51).

Laboratory energy Mev	$-(e^2/\pi)b(0°)/U$ percent	$-(e^2/\pi)b(90°, \text{c.m.})/U$ percent
50	3.80	−0.32
150	5.26	−0.87
300	6.41	−2.13
1000	8.80	−4.35

A few results for high energies computed from (49) and (51) are given in Table II. At 180°, the quantity $-(e^2/\pi)b/U$ is +0.98 percent.

VI. NONRELATIVISTIC LIMIT

In the nonrelativistic or Thompson limit our results will be equally valid in the laboratory and c.m. systems, since we will keep only the first nonvanishing terms. In the c.m. system, we let the scattering angle be φ and $|\mathbf{q}_1| = |\mathbf{q}_2| = \omega$. Then

$$\kappa = -2p_1\cdot q_1 = -2[(1+\omega^2)^{\frac{1}{2}}\omega + \omega^2]$$
$$\tau = 2p_1\cdot q_2 = 2[(1+\omega^2)^{\frac{1}{2}}\omega + \omega^2\cos\varphi], \quad (52)$$

so that for $\omega \ll 1$:

$$\kappa = -2\omega(1+\omega+\omega^2/2+\cdots)$$
$$\tau = 2\omega(1+\omega\cos\varphi+\omega^2/2+\cdots). \quad (53)$$

Since κ depends only on ω and $d\tau = \omega^2 d\Omega/\pi$, Eq. (11) becomes

$$d\sigma = (e^4/2m^2)d\Omega(1-2\omega+2\omega^2+\cdots)[U - (e^2/\pi)U^{(1)}]. \quad (54)$$

In Table III we give the nonrelativistic limits of functions occurring in $U^{(1)}$ (Eq. 30). A straightforward calculation then yields:

$$U = 1+\cos^2\varphi+0(\omega) \quad (55)$$

$$U^{(1)} = -(4/3)\omega^2(1-\cos\varphi)U\ln\lambda + (1+\cos\varphi+\cos^2\varphi-\tfrac{4}{3}\cos^3\varphi)4\omega^2\ln\omega + 0(\omega^2). \quad (56)$$

The angular dependence of the $4\omega^2\ln\omega$ term is plotted as $f(\varphi)$ in Fig. 6. Expression (56) disagrees with the result of Schafroth.[3]

Since $(1/\kappa)+(1/\tau) = \frac{1}{2}(1-\cos\varphi) - (\omega/2)\sin^2\varphi$ in the c.m. system and the same invariant is $\frac{1}{2}(1-\cos\theta)$ in the laboratory system, with θ the laboratory scattering angle, it is clear that $\cos\theta = \cos\varphi + 0(\omega)$ for given κ, τ. Also, since $\kappa+\tau = -2\omega_1\omega_2(1-\cos\theta) = -2\omega^2(1-\cos\varphi)$ and $\omega_1 = \omega_2 + 0(\omega_1^2)$, we get that $\omega_1 = \omega + 0(\omega^2) \approx \omega_2$ and therefore (56) is valid with φ interpreted as the laboratory scattering angle and ω the incident photon energy. Equation (55) is valid, with this interpretation, to order ω^2.

The double Compton effect (39) gives in this limit (with U_D defined as in (47))

$$U_D = -(4/3)\omega^2(1-\cos\varphi)U\ln(2k_{\max}/\lambda) + 0(\omega^2). \quad (57)$$

It will be observed that all the corrections vanish in the zero energy limit.

APPENDIX W. ON THE TRANSCENDENTAL FUNCTIONS $G_0(\kappa)$ AND $h(y)$

The complete expression for the radiative correction (29) is expressed in terms of the relatively unfamiliar transcendental integrals $G_0(\kappa)$ and $h(y)$. These can both be expressed, however, in terms of one of the

so-called Spence functions,[14] namely

$$L(x)=\int_0^x \ln(1-u)du/u \tag{A1}$$

which we shall consider briefly.

It is well known that $L(-1)=\pi^2/12$, $L(1)=-\pi^2/6$. If $x<1$,

$$L(x)=-\int_0^x (u+\tfrac{1}{2}u^2+\cdots)du/u$$
$$=-\left(x+\frac{x^2}{2^2}+\frac{x^3}{3^2}+\cdots\right). \tag{A2}$$

If $x>1$,

$$L(x)=L(1)+\int_1^x \ln(1-u)du/u$$
$$=L(1)+\int_1^x [\ln|1-u|\pm i\pi]du/u$$
$$=\bar{L}(x)\pm i\pi \ln x \tag{A3}$$

where

$$\bar{L}(x)=\int_0^x \ln|1-u|\,du/u.$$

For computational convenience, we can also write for $x>1$:

$$\bar{L}(x)=L(1)+\int_{1/x}^1 \ln\left(\frac{1-v}{v}\right)\frac{dv}{v}$$
$$=-\tfrac{1}{3}\pi^2-L(1/x)+\tfrac{1}{2}(\ln x)^2, \tag{A4}$$

and in a similar manner

$$L(-x)=(\pi^2/6)-L(-1/x)+\tfrac{1}{2}(\ln x)^2. \tag{A5}$$

Since $\kappa=(m^2-p_3^2)$ is always negative, $G_0(\kappa)=2\kappa^{-1}[L(1-\kappa)-L(1)]$ has an imaginary part whose sign is not determined by (A3). To fix the sign we must recall that according to the scheme of reference 1(b), all photons and electrons are considered to have a small additional negative imaginary mass. Thus κ has a small positive imaginary part, i.e., $\kappa=-|\kappa|+i\delta$ with δ vanishingly small. Therefore

$$G_0(\kappa)=2\kappa^{-1}[\bar{L}(1-\kappa)-L(1)]+i\pi 2\kappa^{-1}\ln(1-\kappa). \tag{A6}$$

(Similarly, the term $\ln\kappa$ in (30) is equal to $\ln|\kappa|+i\pi$.) We might also note that since τ is always positive, $G_0(\tau)$ has no imaginary part.

Thus if $-\kappa\gg 1$, $\tau\gg 1$ (from A4, A5):

$$G_0(\kappa)\simeq 2\kappa^{-1}[\tfrac{1}{2}(\ln|\kappa|)^2-(\pi^2/6)+i\pi\ln|\kappa|] \tag{A7}$$

$$G_0(\tau)\simeq 2\tau^{-1}[\tfrac{1}{2}(\ln\tau)^2+(\pi^2/3)]. \tag{A8}$$

[14] For references see Fletcher, Miller, and Rosenhead, *An Index of Mathematical Tables* (Scientific Computing Service, Ltd., London, 1946), p. 343.

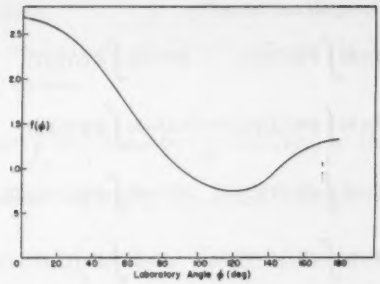


FIG. 6. Angular dependence of nonrelativistic limit of $U^{(1)}$ (without $\ln\lambda$ term).

Our other transcendental function

$$h(y)=y^{-1}\int_0^y u\,du\ \mathrm{ctnh}u$$

can also be expressed in terms of $L(x)$. Integration by parts gives

$$h(y)=\ln(\sinh y)-y^{-1}\int_0^y \ln(\sinh u)du. \tag{A9}$$

Letting $t=e^{-2u}$ yields $\ln(\sinh u)=\ln(1-t)-\ln 2t$ and the integral is obtained directly:

$$h(y)=\ln(2\sinh y)-y/2+(2y)^{-1}[\pi^2/6+L(e^{-2y})]. \tag{A10}$$

If y is sufficiently large that we can neglect e^{-2y} compared to unity, we get

$$h(y)\simeq(y/2)+(\pi^2/12y). \tag{A11}$$

APPENDIX X. TABLE OF INTEGRALS

In this appendix we will simply list the integrals which enter this problem, reserving for Appendix Y a discussion of the methods used. To simplify the presentation of the integrals (which occur also in other problems) we introduce the following definitions:

For factors of the denominator we write $k^2-2p_1\cdot k=(1)$, $k^2-2p_2\cdot k=(2)$, $k^2-2p_3\cdot k-\kappa=(\kappa)$, $k^2=(0)$. We write for frequently occurring vectors

$$p_3=p_1+q_1, \qquad p_4=p_1-q_3,$$
$$2p_0=p_1+p_2, \qquad 2q_0=q_1+q_2,$$
$$2Q=p_1-p_2=q_2-q_1.$$

TABLE III. Nonrelativistic limits of functions occurring in Eq. (30). These expressions are valid in either the c.m. or laboratory system with φ the scattering angle and ω the incident photon energy.

$$\gamma^2=-\tfrac{1}{4}(\kappa+\tau)-(1/48)(\kappa+\tau)^2+0(\omega^6)$$
$$h(y)=1+y^2/9+\cdots$$
$$G_0(\kappa)=2(1+\kappa/2+\kappa^2/9+\cdots)-2\ln|\kappa|(1+\kappa/2+\kappa^2/3+\cdots)$$
$$\ln|\kappa|=\ln 2\omega+\omega+0(\omega^2)$$
$$\ln\tau=\ln 2\omega+\omega\cos\varphi+\tfrac{1}{2}\omega^2(1-\cos^2\varphi)+0(\omega^3)$$

The integrals are defined by

$$A=8i\int d^4k/(1)(2), \qquad B^{(1)}=8i\int d^4k/(1)(\kappa)$$

$$C^{(1)}=8i\int d^4k/(1)(0), \qquad D=8i\int d^4k/(0)(\kappa)$$

$$F=8i\int d^4k/(1)(2)(\kappa), \quad G^{(1)}=8i\int d^4k/(1)(\kappa)(0)$$

$$H=8i\int d^4k/(1)(2)(0), \qquad J=8i\int d^4k/(1)(2)(\kappa)(0).$$

Integrals $B^{(2)}$, $C^{(2)}$, $G^{(2)}$ are defined as $B^{(1)}$, $C^{(1)}$, $G^{(1)}$ but with (2) replacing (1). Their values are obtained from those of $B^{(1)}$, $C^{(1)}$, $G^{(1)}$ by replacing p_1 by p_2, q_1 by q_2, and Q by $-Q$. We use the notation, as in reference 1(b), that

$$F_{(0;\sigma;\sigma\tau)}=8i\int (1;k_\sigma;k_\sigma k_\tau)d^4k/(1)(2)(\kappa), \text{ etc.}$$

Let:

$$\mu=Q^2=-\sinh^2 y=\tfrac{1}{4}(\kappa+\tau), \quad (y \text{ is real as } Q^2<0)$$

$$a=-\ln\Lambda^2$$

$$b=1-y\coth y$$

$$c=(\kappa/\kappa-1)\ln\kappa$$

$$d=2y\,\mathrm{csch}2y=(1-b)/(1-\mu)$$

$$L(x)=\int_0^x \frac{du}{u}\ln(1-u)$$

$$h(y)=(1/y)\int_0^y u\coth u\,du$$

$$v=\kappa^2+4\mu(1-\kappa).$$

Integrals

$$A_0=2a-4b+2$$

$$B_0^{(1)}=2a+2=B_0$$

$$C_0^{(1)}=2a-2=C_0$$

$$D_0=2a-2+2c$$

$$F_0=y^2\,\mathrm{csch}^2 y$$

$$G_0^{(1)}=(2/\kappa)[L(1-\kappa)-L(1)]=G_0$$

$$H_0=2d[-\ln\lambda+h(2y)-h(y)]$$

$$J_0=(2d/\kappa)[2h(y)-h(2y)-\ln(\kappa/\lambda)]$$

$$A_\sigma=(2a-4b+3)p_{0\sigma}$$

$$B_\sigma^{(1)}=(a+\tfrac{3}{2})(p_{1\sigma}+p_{2\sigma})$$

$$C_\sigma^{(1)}=(a+\tfrac{1}{2})p_{1\sigma}$$

$$D_\sigma=[(\kappa/\kappa-1)(c-1)+a+\tfrac{1}{2}]p_{2\sigma}$$

$$F_\sigma=F_0p_{0\sigma}+[F_0-(2b/\mu)]q_{0\sigma}$$

$$G_\sigma^{(1)}=[G_0-(2c/\kappa)]p_{1\sigma}+(2/\kappa)[G_0-(2-\kappa/\kappa)c-2]q_{1\sigma}$$

$$H_\sigma=2dp_{0\sigma}$$

$$vJ_\sigma=[2\mu F_0-(2\mu-\kappa)Z-\kappa G_0]p_{0\sigma}+[(2\mu-\kappa)F_0+2(1-\mu)Z-(2-\kappa)G_0]q_{0\sigma}$$

with

$$Z=H_0+\kappa J_0=2d[-\ln\kappa+h(y)]$$

$$F_{\sigma\tau}=F_\sigma p_{0\tau}+[(F_0-2b/\mu)p_{0\sigma}+(F_0+Y_0-2b/\mu)q_{0\sigma}]q_{0\tau}-Y_0Q_\sigma Q_\tau+[\mu Y_0+\tfrac{1}{2}a+\tfrac{3}{4}]\delta_{\sigma\tau}$$

with

$$Y_0=(F_0-2b-1)/2\mu$$

$$G_{\sigma\tau}^{(1)}=p_{1\sigma}p_{1\tau}\left[G_0-\frac{3\kappa-2}{\kappa-1}\frac{c}{\kappa}+\frac{1}{\kappa-1}\right]+(p_{1\sigma}q_{1\tau}+q_{1\sigma}p_{1\tau})\left[\frac{3}{\kappa}G_0+\frac{2\kappa^2-9\kappa+6}{\kappa(\kappa-1)}\frac{c}{\kappa}+\frac{6-5\kappa}{\kappa-1}\frac{1}{\kappa}\right]+q_{1\sigma}q_{1\tau}\left[\frac{6}{\kappa^2}G_0+\frac{\kappa^3+4\kappa^2-18\kappa+12}{\kappa^2(\kappa-1)}\frac{c}{\kappa}-\frac{2\kappa^2+9\kappa-12}{\kappa(\kappa-1)}\frac{1}{\kappa}\right]+\frac{1}{4}\delta_{\sigma\tau}\left[2G_0-3+2a+2\frac{\kappa-2}{\kappa}c\right]$$

$$H_{\sigma\tau}=dp_{0\sigma}p_{0\tau}+(b/\mu)Q_\sigma Q_\tau+\tfrac{1}{2}\delta_{\sigma\tau}(a-2b+\tfrac{1}{2})$$

$$J_{\sigma\tau}=\alpha_\sigma Q_\tau+\beta_\sigma p_{0\tau}+\gamma_\sigma p_{3\tau}+\epsilon s_{\sigma\tau}$$

where:

$$\epsilon=F_0-\alpha_\sigma Q_\sigma-\beta_\sigma p_{0\sigma}-\gamma_\sigma P_{3\sigma}$$

$$4\mu\alpha_\sigma=G_\sigma^{(1)}-G_\sigma^{(2)}$$

$$v\beta_\sigma=\kappa F_\sigma-(2-\kappa)(H_\sigma+\kappa J_\sigma)+(1-\kappa)(G_\sigma^{(1)}+G_\sigma^{(2)})$$

$$v\gamma_\sigma=(2\mu-\kappa)F_\sigma+2(1-\mu)(H_\sigma+\kappa J_\sigma)-\tfrac{1}{2}(2-\kappa)(G_\sigma^{(1)}+G_\sigma^{(2)})$$

$$s_{\sigma\tau}=\delta_{\sigma\tau}-(Q_\sigma Q_\tau/\mu)+4(1-\kappa)(p_{0\sigma}p_{0\tau}/v)-2(2-\kappa)[(p_{0\sigma}p_{3\tau}+p_{3\sigma}p_{0\tau})/v]+4(1-\mu)(p_{3\sigma}p_{3\tau}/v)$$

$$[s_{\sigma\sigma}=1]$$

$$J_{\sigma\tau\rho}=\alpha_{\sigma\tau}Q_\rho+\beta_{\sigma\tau}p_{0\rho}+\gamma_{\sigma\tau}p_{3\rho}+\epsilon_\sigma s_{\tau\rho}+\epsilon_\tau s_{\sigma\rho}$$

where

$$\epsilon_\sigma=F_\sigma-\alpha_{\sigma\tau}Q_\tau-\beta_{\sigma\tau}p_{0\tau}-\gamma_{\sigma\tau}p_{3\tau}$$

$$4\mu\alpha_{\sigma\tau}=G_{\sigma\tau}^{(1)}-G_{\sigma\tau}^{(2)}$$

$$v\beta_{\sigma\tau}=\kappa F_{\sigma\tau}-(2-\kappa)(H_{\sigma\tau}+\kappa J_{\sigma\tau})+(1-\kappa)(G_{\sigma\tau}^{(1)}+G_{\sigma\tau}^{(2)})$$

$$v\gamma_{\sigma\tau}=(2\mu-\kappa)F_{\sigma\tau}-2(1-\mu)(H_{\sigma\tau}+\kappa J_{\sigma\tau})-\tfrac{1}{2}(2-\kappa)(G_{\sigma\tau}^{(1)}+G_{\sigma\tau}^{(2)}).$$

APPENDIX Y. ON THE INTEGRALS IN THE CORRECTED CROSS SECTION

In this section we will describe the methods by which the integrations occurring in the e^4-order matrix element have been performed and will give some examples of the calculations.

Those integrals which are scalars (those with no k in the numerator) have been done by the parametric method discussed in the appendix to reference 1(b). Those which are tensors (having one or more k's in the numerator) can be done either by the parametric method, or can be derived by an algebraic procedure described below from those of lower tensor order.

We now do several examples of the integrals to indicate the methods employed.

(a) The Two-Denominator Integrals

These all have the form $\chi(\Lambda^2)$ (where we need only the case of Λ much greater than any of the momenta involved in the problem), with

$$\chi=8i\int(1;k_\sigma)(k^2-2p_1\cdot k-\Delta_1)^{-1}\times(k^2-2p_2\cdot k-\Delta_2)^{-1}(-\Lambda^2)(k^2-\Lambda^2)^{-1}d^4k. \quad \text{(A12)}$$

To reduce this integral and those considered below we use the methods of reference 1(b).

In the first place we combine the denominators by making use of the relation

$$1/ab=\int_0^1 dy[ay+b(1-y)]^{-2} \quad \text{(A13)}$$

and similar expressions for $1/ab^2$, $1/ab^3$, etc., obtainable from (A13) by differentiation. Thus χ becomes

$$\chi=8i\int_0^1\int_0^1\int(1;k_\sigma)dy2zdz(-\Lambda^2)d^4k\times[k^2-2zp_y\cdot k-z\Delta_y-(1-z)\Lambda^2]^{-3} \quad \text{(A14)}$$

with $p_y=p_1y+p_2(1-y)$; $\Delta_y=\Delta_1y+\Delta_2(1-y)$. Using (12a) of reference 1(b), we get

$$\chi=\int_0^1\int_0^1 2zdzdy(-\Lambda^2)(1;zp_{y\sigma})\times[z^2p_y^2+z\Delta_y+(1-z)\Lambda^2]^{-1}. \quad \text{(A15)}$$

To facilitate the work we observe that in the limit of very large Λ^2 (with $\delta\ll1$)

$$\int_0^1\frac{(1-z)^n dz(-\Lambda^2)}{P(z)+\Lambda^2(1-z)}\cong-\int_0^{1-\delta}(1-z)^{n-1}dz+\int_{1-\delta}^1\frac{(1-z)^n dz(-\Lambda^2)}{P(1)+\Lambda^2(1-z)}\cong-1/n,\quad n>0\cong\ln[P(1)/\Lambda^2],\quad n=0. \quad \text{(A16)}$$

Writing

$$(1;zp_{y\sigma})=(1-(1-z);[(1-z)^2-2(1-z)+1]p_{y\sigma}),$$

(A15) becomes

$$\chi=\int_0^1 2dy\left\{(1;p_{y\sigma})\ln\frac{p_y^2+\Delta_y}{\Lambda^2}+(1;\tfrac{2}{3}p_{y\sigma})\right\}. \quad \text{(A17)}$$

The second term gives $(2;3(p_{1\sigma}+p_{2\sigma}))$. In the first term

$$\chi_a=\int_0^1 2dy(1;p_{y\sigma})\ln[(p_y^2+\Delta_y)/\Lambda^2], \quad \text{(A18)}$$

notice that with $2Q=(p_1-p_2)$,

$$p_y^2=p_2^2+Q^2y^2+2p_2\cdot(p_1-p_2)y,\qquad \Delta_1=m^2-p_1^2,\quad \Delta_2=m^2-p_2^2$$

so that

$$p_y^2+\Delta_y=m^2+4Q^2(y^2-y)=m^2+Q^2[(2y-1)^2-1].$$

Now let $m=1$, $Q^2=\sin^2\theta$, and $2y-1=\tan\alpha/\tan\theta$, so that $dy(\sec^2\alpha/2\tan\theta)d\alpha$. We get

$$p_y^2+\Delta_y=\cos^2\theta\sec^2\alpha$$

so that

$$\chi_a=\int_{-\theta}^{\theta}(\sec^2\alpha/\tan\theta)d\alpha(1;p_{y\sigma}(\alpha))\times\ln(\cos^2\theta\sec^2\alpha/\Lambda^2). \quad \text{(A19)}$$

This integral can be done easily. For example,

$$\int_{-\theta}^{\theta}\sec^2\alpha d\alpha\ln(\sec^2\alpha)=\int_{-\tan\theta}^{\tan\theta}dy\ln(1+y^2)=4\tan\theta[\ln(\sec\theta)-1]+4\theta,$$

the last integral being performed by parts. In this manner we obtain integrals A, B, C, D.

(b) The Intregal $G_0^{(1)}$

As an example of the three denominator integrals we integrate $G_0^{(1)}$. The parameterization method gives

$$G_0^{(1)}=8i\int d^4k(k^2-2p_1\cdot k)^{-1}(k^2-2p_2\cdot k-\kappa)^{-1}(k^2)^{-1}=8i\int_0^1\int_0^1\int d^4k2dyxdx(k^2-2p_x\cdot k-\Delta_x)^{-3} \quad \text{(A20)}$$

with

$$p_y=(1-y)p_1+yp_2=p_1+yq_1,\quad p_x=xp_y,\quad \Delta_x=xy\kappa.$$

Using (12a) of reference 1(b) again,

$$G_0^{(1)}=\int_0^1\int_0^1 2dxdy(xp_y^2+y\kappa)^{-1}. \quad \text{(A21)}$$

Since

$$\int_0^1 dx(ax+b)^{-1}=a^{-1}\ln(b+a)/b,$$

and since $p_y^2=1-\kappa y$

$$G_0^{(1)}=-\int_0^1 2dy(1-\kappa y)^{-2}\ln\kappa y$$

$$=-\frac{2}{\kappa}\int_{1-\kappa}^1\frac{dv}{v}\ln(1-v). \quad \text{(A22)}$$

In the last step we have let $v=1-\kappa y$. We obtain finally:

$$G_0^{(1)}=G_0^{(2)}=G_0=(2/\kappa)[L(1-\kappa)-L(1)]. \quad \text{(A23)}$$

(c) The Integrals H

These are the same as those done in the radiationless scattering problem. They are given in reference 1(b), appendix. Equations (23a), (24a), and (25a) should have the signs of their left-hand sides changed.

(d) The Integral J_0

$$J_0=8i\int d^4k(k^2-2p_1\cdot k)^{-1}(k^2-2p_2\cdot k-\kappa)^{-1}$$

$$\times(k^2-2p_2\cdot k)^{-1}(k^2-\lambda^2)^{-1}. \quad \text{(A24)}$$

We have given the photon k a small mass λ as an infrared cutoff. When this integral is parameterized it becomes

$$J_0=8i\int_0^1\int_0^1\int_0^1\int 6dxdydz(1-x)z^2d_4k$$

$$\times[k^2-2zp_z\cdot k-xz\kappa-\lambda^2(1-z)]^{-4}. \quad \text{(A25)}$$

By (13a) of reference 1(b),

$$8i\int d^4k(k^2-2p\cdot k-\Delta)^{-4}=-\tfrac{1}{3}(p^2+\Delta)^{-2}$$

so that

$$J_0=-\int_0^1\int_0^1\int_0^1 2dxdydz(1-x)z^2$$

$$\times[z^2p_z^2+xz\kappa+\lambda^2(1-z)]^{-2} \quad \text{(A26)}$$

with $p_z=(1-x)p_y+xp_2$, $p_y=yp_1+(1-y)p_2$.

We break the x integration into two regions ($\epsilon\ll 1$):

$$\int_0^1 dx\int_0^1 dy\int_0^1 dz=\int_0^1 dy\int_\epsilon^1 dx\int_0^1 dz+\int_0^1 dy\int_0^\epsilon dx\int_0^1 dz.$$

(I) (II)

In (I) we let $\lambda\to 0$, getting

$$(\mathrm{I})=\int_0^1 dy\int_\epsilon^1 dx\int_0^1 2dz(1-x)(zp_z^2+x\kappa)^{-2}$$

$$=2\int_0^1\int_\epsilon^1\frac{(1-x)dxdy}{\kappa x(p_z^2+\kappa x)}. \quad \text{(A27)}$$

In region (II), since x is small, we neglect x^2 compared to x:

$$(\mathrm{II})=\int_0^1 dy\int_0^1 dz\int_0^\epsilon 2z^2dx[z^2p_y^2+2xz^2a$$

$$+z\kappa x+\lambda^2(1-z)]^{-2}$$

$$=2\int_0^1 dy\int_0^1 dz\epsilon z^2[z^2p_y^2+\lambda^2(1-z)]^{-1}$$

$$\times[z^2p_y^2+2\epsilon z^2a+z\kappa\epsilon+\lambda^2(1-z)]^{-1} \quad \text{(A28)}$$

with $a=p_y\cdot(p_2-p_y)$.

In (II) we now break the z integration into two regions; $0\le z\le z_c$ and $z_c\le z\le 1$ such that $\lambda^2\ll z_c^2p_y^2\ll z_c\kappa\epsilon$. Thus for $z<z_c$ we neglect z relative to unity and $z^2p_y^2$ relative to $z\kappa\epsilon$. For $z>z_c$ we neglect λ. There results:

$$\int_{z_c}^1\frac{2\epsilon dzdy}{p_y^2z[z(p_y^2+2\epsilon a)+\kappa\epsilon]}=\frac{2}{\kappa p_y^2}\ln\frac{\kappa\epsilon}{z_cp_y^2},$$

$$\int_0^{z_c}\frac{2\epsilon z^2dzdy}{(z^2p_y^2+\lambda^2)(z\kappa\epsilon+\lambda^2)}$$

$$=-\int_0^{z_c}\frac{2\epsilon dz}{(\lambda^2p_y^2+\kappa^2\epsilon^2)}\left(\frac{\lambda^2-\kappa\epsilon z}{z^2p_y^2+\lambda^2}-\frac{\lambda^2}{\kappa\epsilon z+\lambda^2}\right)$$

$$=\frac{1}{\kappa p_y^2}\ln\frac{z_c^2p_y^2}{\lambda^2}(\text{neglecting terms of order }\lambda).$$

Adding these together we get

$$(\mathrm{II})=\int_0^1\frac{dy}{\kappa p_y^2}\ln\frac{\kappa^2\epsilon^2}{\lambda^2p_y^2}. \quad \text{(A29)}$$

We can now use the same substitution for y as that leading to (A8). Notice that in this case $\Delta_y=0$. We get $dy/\kappa p_y^2=d\alpha/\sin 2\theta$, $p_y^2=\cos^2\theta/\cos^2\alpha$ so that (II) becomes:

$$(\mathrm{II})=\int_{-\theta}^{\theta}\frac{2d\alpha}{\sin 2\theta}\left[\ln\frac{\kappa\epsilon}{\lambda\cos\theta}+\ln(\cos\alpha)\right]$$

$$=\frac{4\theta}{\sin 2\theta}\ln\frac{\kappa\epsilon}{\lambda\cos\theta}+\int_0^\theta\frac{4d\alpha}{\sin 2\theta}\ln(\cos\alpha). \quad \text{(A30)}$$

We have still to finish evaluating (I), Eq. (A27). First investigate the denominator $p_x^2+\kappa x$. Since $p_3^2=1-\kappa$ and $2p_3\cdot p_y=2-\kappa$,

$$p_x^2=(1-x)^2p_y^2+x^2(1-\kappa)+x(1-x)(2-\kappa)$$
$$p_x^2+\kappa x=(1-x)^2(p_y^2-1)+1.$$

Letting $1-x=\sin\phi/\sin\theta$, $2y-1=\tan\alpha/\tan\phi$ with $Q^2=\sin^2\theta$, we get

$$p_y^2=\sin^2\theta[(\tan^2\alpha/\tan^2\phi)-1],$$
$$p_x^2+\kappa x=\sin^2\phi\,\mathrm{ctn}^2\phi(\sec^2\alpha-\sec^2\phi)+1=\cos^2\phi\,\sec^2\phi,$$
$$-dx=\cos\phi d\phi/\sin\phi,\quad dy=\sec^2\alpha d\alpha/2\tan\phi.$$

The integrand of (I) becomes

$$\frac{1}{\kappa}\left(\frac{\sin\theta}{\sin\theta-\sin\phi}\right)\left(\frac{\sin\phi}{\sin\theta}\right)\left(\frac{\cos^2\alpha}{\cos^2\phi}\right)\left(\frac{\cos\phi d\phi}{\sin\theta}\right)\left(\frac{\sec^2\alpha d\alpha}{\tan\phi}\right)$$
$$=\frac{1}{\kappa}\frac{d\phi d\alpha}{\sin\theta-\sin\phi}$$

and

$$(\mathrm{I})=\int_0^{\theta-\epsilon\tan\theta}\frac{d\phi}{\kappa}\int_{-\phi}^{\phi}\frac{d\alpha}{\sin\theta(\sin\theta-\sin\phi)}$$
$$=\int_0^{\theta-\epsilon\tan\theta}\frac{2\phi d\phi}{\kappa\sin\theta(\sin\theta-\sin\phi)}\qquad(\mathrm{A31})$$

Integrate (A31) by parts, using No. 436, Dwight's *Tables of Integrals*, to get

$$(\mathrm{I})=-\frac{4\theta}{\kappa\sin2\theta}\ln\left(\frac{\epsilon\tan\theta}{2\cos\theta}\right)+\int_0^{\theta}\frac{4d\phi}{\sin2\theta}\ln\left(\frac{\sin\frac{1}{2}(\theta-\phi)}{\cos\frac{1}{2}(\theta+\phi)}\right)\qquad(\mathrm{A32})$$

which can be put in the form

$$(\mathrm{I})=-\frac{4\theta}{\kappa\sin2\theta}\ln\frac{\epsilon\sin\theta}{\cos^2\theta}-\int_0^{\theta}\frac{4d\alpha}{\sin2\theta}\ln(\cos\alpha)+\int_0^{\theta}\frac{4d\alpha}{\sin2\theta}\ln(\tan\alpha)$$

and can now be added to (A19).

This gives the result,

$$J_0=\frac{-4\theta}{\kappa\sin2\theta}\left[\ln\frac{\kappa}{\lambda\tan\theta}+\theta^{-1}\int_v^{\theta}dv\ln(\tan v)\right].\qquad(\mathrm{A33})$$

Another integration by parts and the substitution $\theta=iy$ gives the result in the table:

$$J_0=(2d/\kappa)[2h(y)-h(2y)-\ln(\kappa/\lambda)].\qquad(\mathrm{A34})$$

(e) The Integrals J_σ, $J_{\sigma\tau}$, $J_{\sigma\tau\nu}$

From the preceding work on J_0 it may be supposed that to attempt these more complicated integrals by the parametric method would involve great labor. Fortunately there is a way to reduce these integrals to a combination of integrals of a lower tensor order, and those of a smaller number of denominators. We will do J_σ as an example but it should be clear from this how $J_{\sigma\tau}$ and $J_{\sigma\tau\nu}$ are done. This method can of course, be applied also to G_σ, $G_{\sigma\tau}$, etc. We shall be able to express J_σ in terms of J_0 and the integrals F, G having only three denominators.

Using the notation indicated in Appendix X for the denominators, $(1)=k^2-2p_1\cdot k$, etc., we write

$$J_\sigma=8i\int\frac{k_\sigma d^4k}{(1)(2)(\kappa)(0)}=\alpha p_{1\sigma}+\beta p_{2\sigma}+\gamma p_{3\sigma},\qquad(\mathrm{A35})$$

α, β, γ being scalar functions of p_1, p_2, and p_3. The vectors p_1, p_2, p_3 will in general define a three-space. It is clear that the vector J_σ cannot have a component in the direction P which is perpendicular to this three-space, since for k_σ in the P direction, the integrand is an odd function and therefore J_σ must vanish.

If we now take the scalar products of J_σ with p_1, p_2, p_3, since $2p_1\cdot k=(0)-(1)$, etc., we get

$$2p_1\cdot J=8i\int\frac{2p_1\cdot kd^4k}{(1)(2)(\kappa)(0)}$$
$$=8i\int\frac{d^4k}{(1)(2)(\kappa)}-8i\int\frac{d^4k}{(\kappa)(2)(0)}=F_0-G_0$$
$$2p_2\cdot J=8i\int\frac{2p_2\cdot kd^4k}{(1)(2)(\kappa)(0)}$$
$$=8i\int\frac{d^4k}{(1)(2)(\kappa)}-8i\int\frac{d^4k}{(\kappa)(2)(0)}=F_0-G_0\qquad(\mathrm{A36})$$
$$2p_3\cdot J=8i\int\frac{2p_3\cdot kd^4k}{(1)(2)(\kappa)(0)}$$
$$=8i\int\frac{d^4k}{(1)(2)(\kappa)}-8i\int\frac{d^4k}{(1)(2)(0)}-8i\int\frac{\kappa d^4k}{(1)(2)(\kappa)(0)}$$
$$=F_0-H_0-\kappa J_0.$$

Taking also the scalar products with the right-hand side of (A35), we get the set of linear equations:

$$F_0-G_0=2\alpha+\beta(1-\tfrac{1}{2}\kappa-\tfrac{1}{2}\tau)+\gamma(2-\kappa)$$
$$F_0-G_0=\alpha(1-\tfrac{1}{2}\kappa-\tfrac{1}{2}\tau)+2\beta+\gamma(2-\kappa)\qquad(\mathrm{A37})$$
$$F_0-H_0-\kappa J_0=(\alpha+\beta)(2-\kappa)+2\gamma(1-\kappa).$$

J_σ can now be readily obtained by solving these equations for α, β, and γ.

To obtain $J_{\sigma\tau}$ we write

$$J_{\sigma\tau}=8i\int\frac{k_\sigma k_\tau d^4k}{(1)(2)(\kappa)(0)}$$
$$=\alpha_\sigma p_{1\tau}+\beta_\sigma p_{2\tau}+\gamma_\sigma p_{3\tau}+\epsilon\delta_{\sigma\tau}, \quad \text{(A38)}$$

α_σ, β_σ, γ_σ being vector functions of p_1, p_2, p_3, and ϵ a scalar function of the same variables. The tensor $\epsilon\delta_{\sigma\tau}$ now occurs on the right-hand side as it is possible for $J_{\sigma\tau}$ to have nonzero components depending on P. (If $k_\sigma=k_\tau=k_P$, $J_{\sigma\tau}$ need not vanish.) If we take inner products with p_1, p_2, and p_3 now we get

$$F_\sigma-G_\sigma^{(1)}=2\alpha_\sigma+2\beta_\sigma(1-\tfrac{1}{2}\kappa-\tfrac{1}{2}\tau)+\gamma_\sigma(2-\kappa)+2\epsilon p_{1\sigma}$$

for $2p_{1\tau}J_{\sigma\tau}$ and similar equations for p_2 and p_3. This gives us three equations for the four quantities α_σ, β_σ, γ_σ, ϵ. However, there is the additional independent result obtained by summing $J_{\sigma\sigma}$ over σ:

$$J_{\sigma\sigma}=8i\int\frac{k_\sigma k_\sigma d^4k}{(1)(2)(\kappa)(0)}$$
$$=F_0=\alpha_\sigma p_{1\sigma}+\beta_\sigma p_{2\sigma}+\gamma_\sigma p_{3\sigma}+4\epsilon. \quad \text{(A39)}$$

Solving these four equations we obtain $J_{\sigma\tau}$ algebraically in terms of simpler integrals. In a similar manner $J_{\sigma\tau\rho}$ can be expressed in the form given in Appendix X.

APPENDIX Z. EXAMPLES OF THE CALCULATIONS

We shall here illustrate by two examples the method of evaluation of the transition amplitude.

The matrix T (18) can be written, if we rationalize the denominator, as ($p^2=m^2$, $q^2=0$)

$$T=\int(k^2-2p\cdot k-2q\cdot k-\kappa)^{-1}(k^2-2p\cdot k)^{-1}$$
$$\times d^4k\boldsymbol{k}^{-2}C(\boldsymbol{k}^2)\boldsymbol{T} \quad \text{(A40)}$$

with

$$\boldsymbol{T}=\gamma_\mu(\boldsymbol{p}+\boldsymbol{q}-\boldsymbol{k}+m)\boldsymbol{e}(\boldsymbol{p}-\boldsymbol{k}+m)\gamma_\mu. \quad \text{(A41)}$$

$\boldsymbol{T}$ can now be split into terms involving no $\boldsymbol{k}$, one $\boldsymbol{k}$, and two $\boldsymbol{k}$'s, and the results of the integrations over $\boldsymbol{k}$-space inserted from Appendix Y. Thus,

$$\boldsymbol{T}=\gamma_\mu(\boldsymbol{p}+\boldsymbol{q}+m)\boldsymbol{e}(\boldsymbol{p}+m)\gamma_\mu-\gamma_\mu\boldsymbol{k}\boldsymbol{e}(\boldsymbol{p}+m)\gamma_\mu$$
$$-\gamma_\mu(\boldsymbol{p}+\boldsymbol{q}+m)\boldsymbol{e}\boldsymbol{k}\gamma_\mu+\gamma_\mu\boldsymbol{k}\boldsymbol{e}\boldsymbol{k}\gamma_\mu$$
$$=2\boldsymbol{p}(\boldsymbol{p}+\boldsymbol{q}+m)\boldsymbol{e}-2\boldsymbol{p}\boldsymbol{k}\boldsymbol{e}+2\boldsymbol{k}\boldsymbol{e}(\boldsymbol{p}+\boldsymbol{q})$$
$$-4m(e\cdot k)-2\boldsymbol{e}\boldsymbol{k}\boldsymbol{k},$$

where we have used Eq. (4a) of reference 1(b) and the fact that the matrix T operates from the left on a state u such that $\boldsymbol{p}u=mu$. Inserting the integrals and grouping terms we now have,

$$8iT=2\boldsymbol{p}(\boldsymbol{p}+\boldsymbol{q}+m)\boldsymbol{e}G_0-2(\boldsymbol{p}\gamma_\sigma\boldsymbol{e}-\gamma_\sigma\boldsymbol{e}(\boldsymbol{p}+\boldsymbol{q})+2me_\sigma)G_\sigma$$
$$-2\gamma_\sigma\boldsymbol{e}\gamma_\tau G_{\sigma\tau}. \quad \text{(A42)}$$

This expression can now be further simplified. For example, the term in G_σ is (using Appendix X)

$$-2[G_0-(2c/\kappa)][\boldsymbol{p}^2\boldsymbol{e}-\boldsymbol{p}\boldsymbol{e}(\boldsymbol{p}+\boldsymbol{q})+2m(e\cdot p)]$$
$$-(4/\kappa)[G_0-(2-\kappa/\kappa)c-2][\boldsymbol{p}\boldsymbol{q}\boldsymbol{e}-\boldsymbol{q}\boldsymbol{e}(\boldsymbol{p}+\boldsymbol{q})]$$

with $\kappa=-2p\cdot q$. Since $\boldsymbol{p}^2=m^2$, $\boldsymbol{q}^2=0$, $e\cdot q=0$, and quite generally $\boldsymbol{a}\boldsymbol{b}+\boldsymbol{b}\boldsymbol{a}=2a\cdot b$, this term becomes finally

$$-2[G_0-(2c/\kappa)](2m^2\boldsymbol{e}-\boldsymbol{p}\boldsymbol{e}\boldsymbol{q})$$
$$+(4/\kappa)[G_0-(2-\kappa/\kappa)c-2]\cdot[2(e\cdot p)\boldsymbol{q}+\kappa\boldsymbol{e}].$$

Combining this with the terms in G_0 and $G_{\sigma\tau}$ in (A42) (expanded in the same manner) we obtain the expression for T given in the text (19).

To illustrate the simplification that occurs upon taking the spur, for unpolarized light, consider the term J (21). This may be decomposed, as was T above, into a sum of terms involving various numbers of $\boldsymbol{k}$'s in the numerator. For example, the term involving three $\boldsymbol{k}$'s is

$$-\int\gamma_\mu\boldsymbol{k}\boldsymbol{e}_2\boldsymbol{k}\boldsymbol{e}_1\boldsymbol{k}\gamma_\mu d^4k/(1)(2)(\kappa)(0) \quad \text{(A43)}$$

using the notation of Appendix X. If we replace $\boldsymbol{e}_1$ by γ_α and $\boldsymbol{e}_2$ by γ_β, this gives a contribution to the spur $P(\kappa,\tau)$, Eq. (28), of a numerical factor times

$$\int g(k)d^4k/(1)(2)(\kappa)(0) \quad \text{(A44)}$$

with

$$g(k)=\mathrm{Sp}[(\boldsymbol{p}_2+m)\gamma_\alpha\boldsymbol{k}\gamma_\beta\boldsymbol{k}\gamma_\alpha\boldsymbol{k}\gamma_\mu(\boldsymbol{p}_1+m)\bar{W}]$$
$$=\kappa^{-1}\{4(p_1\cdot k)(p_2\cdot k)(p_3\cdot k)$$
$$+k^2[\tfrac{1}{2}m^2\kappa(p_1+p_2)\cdot k-(p_1\cdot p_2+2m^2)(p_3\cdot k)]\}$$
$$+\tau^{-1}\{4(p_1\cdot k)(p_2\cdot k)(p_4\cdot k)$$
$$-m^2k^2(p_1+p_2+p_3)\cdot k\}. \quad \text{(A45)}$$

It will now be seen that the factor k^2 will cancel the factor (0) in the denominator of (A44) leading to the integral F_σ (Appendix X). Also, we can write

$$2p_1\cdot k=-(k^2-2p_1\cdot k)+k^2=-(1)+(0)$$

leading to integrals $G_{\sigma\tau}^{(2)}$ and $F_{\sigma\tau}$. Thus it is not necessary, for unpolarized light, to use integral $J_{\sigma\tau\rho}$. (A44) can now be written

$$2(\kappa^{-1}p_{2\sigma}p_{3\tau}+\tau^{-1}p_{2\sigma}p_{4\tau})(F_{\sigma\tau}-G_{\sigma\tau}^{(2)})$$
$$+[\tfrac{1}{2}m^2(p_{1\sigma}+p_{2\sigma})-\kappa^{-1}(p_1\cdot p_2+2m^2)p_{2\sigma}$$
$$-m^2\tau^{-1}(p_{1\sigma}+p_{2\sigma}+p_{3\sigma})]F_\sigma. \quad \text{(A46)}$$

Mean-Value Calculations for Projected Multiple Scattering*

W. T. Scott
Brookhaven National Laboratory, Upton, New York, and Smith College, Northampton, Massachusetts
(Received October 5, 1951)

The multiple-scattering theories of Molière and Snyder-Scott are compared and the equivalence of the mathematical development stated. Using the preferable single-scattering probability of the Molière theory, results are quoted of interest to experimenters. Several mean-value quantities are given: mean arithmetic angle, median angle, half-width, $1/e$ width, angle $1/P_0\pi^{\frac{1}{2}}$ related to the zero-angle amplitude, and mean arithmetic angle with a cutoff at 4 times the mean. These quantities are given for both the projected tangent angle and projected chord angle distributions, in the form of linear relationships between the square of the angle divided by Ω and the logarithm of Ω, where Ω is the mean number of scatterings undergone by the particle in question. The linear relationships are good to 1 percent for Ω from 10^2 to 10^5. Information is also given on smoothed-out distributions, and on an estimate of the error for the cut-off arithmetic mean angle. The scattering constant K is given for several methods of measurement, for Ilford G-5 emulsions.

THE multiple scattering theories of Molière[1] and Snyder and Scott[2,3] are based on the same mathematical development. Molière has an extensive discussion of the single-scattering law used in this development, and his results are superior to those of Snyder and Scott in this regard. It is the purpose of this article to state explicitly the relations between the two theories and to quote some formulas for various types of mean value that have been derived from the numerical tables of references 2 and 3.

We shall use the same notation as given in the previous papers. Molière's single-scattering formula may be obtained from that of Snyder and Scott by (1) replacing the angular unit η_0 by $\chi_\gamma=1.13\eta_0\sqrt{(1.13+3.76\gamma^2)}$ and (2) replacing the mean number of scatterings z by[4] $\Omega=z\eta_0^2/\chi_\gamma^2=z(1.13)^{-2}(1.13+3.76\gamma^2)^{-1}$. Here $\gamma=ZZ'/137\beta$; Z, Z' are, respectively, the atomic numbers of the scattered particle and the scattering nucleus, βc is the velocity of the scattered particle, the factor 1.13 in front of the radical is the ratio of the Bohr radius to the Fermi-Thomas radius for hydrogen, and the numbers 1.13 and 3.76 under the radical sign arise from an empirical fit[1] to a Fermi-Thomas single-scattering calculation. We can thus use our numerically calculated functions[2] $W(\eta|z)$ for the projected angular deviation η after a pathlength corresponding to z, with Molière's corrections, by writing for the distribution in actual angle Φ:

$$P(\Phi)d\Phi=W(\Phi/\chi_\gamma|\Omega)d\Phi/\chi_\gamma. \tag{1}$$

The small-angle approximation is used, and the functions are so normalized that

$$\int_{-\infty}^{\infty}Pd\Phi=1.$$

Molière's calculations depend on the introduction of a quantity B, which is a function of Ω, and a subsequent expansion in inverse powers of B, using as an angular variable the quantity $\varphi=\Phi/\chi_\gamma(B\Omega)^{\frac{1}{2}}$.

For completeness, we give here the formulas for Ω and χ_γ:

$$\Omega=\frac{4\pi NtZ^{4/3}Z'^2\hbar^2}{1.13^2m^2c^2\beta^2(1.13+3.76\gamma^2)}, \tag{2a}$$

$$=7830\sigma Z^{4/3}Z'^2/A\beta^2(1+3.33\gamma^2); \tag{2b}$$

$$\chi_\gamma=1.13mc^2Z^{\frac{1}{2}}(1.13+3.76\gamma^2)^{\frac{1}{2}}/\hbar p \text{ radians}, \tag{3a}$$

$$=0.257Z^{\frac{1}{2}}(1+3.33\gamma^2)^{\frac{1}{2}}/pc \text{ degrees}. \tag{3b}$$

In case the scattering occurs in a mixture of materials, we must find χ_γ and Ω from the separate nuclear species (denoted by index i) by the following:[5]

$$\log\chi_\gamma=\Sigma_i(N_iZ_i^2\log\chi_{\gamma i})/\Sigma_i(N_iZ_i^2) \tag{4}$$

and

$$\chi_\gamma^2\Omega=4\pi te^4Z'^2(pc\beta)^{-2}\Sigma_iN_iZ_i^2 \text{ radians}^2. \tag{5a}$$

$$=515\sigma Z'^2\langle Z^2\rangle_{\text{Av}}/Ap^2c^2\beta^2 \text{ deg}^2. \tag{5b}$$

In these formulas, N is the number of scattering nuclei per unit volume, t is the length of scattered track, σ is the track length measured in grams/cm^2 of scatterer, A is the atomic weight of the scatterer, m and e refer to the electron mass and charge, p is the momentum of the scattered particle, and pc in formulas (3b) and (5b) is measured in Mev/c. The case of energy loss in the scattering material may be included by the method of Molière.[1] The effect of electron-electron scattering can be included for those cases in which it is pertinent (small angles, small Z) by writing $Z_i^2+Z_i$ for Z_i^2 in Eqs. (4) and (5a). This correction has been discussed by Hanson, Lanzl, Lyman, and Scott.[6]

We have compared our tables with the results of the first three terms of Molière's expansion and find excellent numerical agreement for six values of Ω from 100 to 84000—in most cases better than 1 percent and always within about 2 percent. Molière's functions $f^{(1)}$

* Work performed under the auspices of the AEC.
[1] G. Molière, Z. Naturforsch. **2a**, 133 (1947), and **3a**, 78 (1948).
[2] H. S. Snyder and W. T. Scott, Phys. Rev. **76**, 220 (1949).
[3] W. T. Scott and H. S. Snyder, Phys. Rev. **78**, 223 (1950).
[4] Ω is the same as Ω_0 in reference 1.
[5] Reference 1, second paper, Eq. (6.10′).
[6] Hanson, Lanzl, Lyman, and Scott, Phys. Rev. **84**, 634 (1951).

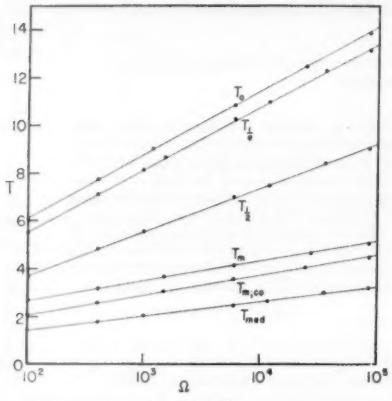


FIG. 1.

and $f^{(2)}$ are not tabulated to much greater accuracy. We have also extended his tables of these functions for φ from 4 to 13.5 in order to compare the tails of our curves, and again have found excellent agreement. The asymptotic formula in reference 2 can be readily shown to agree with that of reference 1; however, the arrangement of terms is different. The expansion given in reference 2 is equivalent to the terms in the same powers of φ in the $f^{(1)}$, $f^{(2)}$, and $f^{(3)}$ terms of reference 1. Terms in higher inverse powers of φ are given by Molière for $f^{(1)}$ and $f^{(2)}$, and it is with these terms that our numerical tables agree. Butler[7] has found an asymptotic development by another method which is almost identical with Molière's $f^{(1)}$ term alone. The $f^{(2)}$ term for $\Omega=100$ contributes 3 percent at $\varphi=4$ and 1 percent at $\varphi=13$. These relative contributions are reduced to one-half at $\Omega=24000$.

The use of numerical tables or graphs, however, is tedious for purposes of comparison with experiment. We have therefore calculated several mean-value expressions that may be of use in various experiments. It turns out that if we plot, for any of these values ψ, the quantity $T_\psi=\psi^2/\chi_\gamma^2\Omega$ as a function of $\log\Omega$, an almost straight line results. Consequently we give our results in the form of such linear relationships, which are sufficiently accurate for most experimental purposes.

The quantities chosen are the following: mean absolute angle, median absolute angle, angle for which $P=P_0/2$ ("half-width"), angle for which $P=P_0/e$ ("$1/e$ width"), the angle $\Phi_0=1/P_0\pi^{\frac{1}{2}}$ derived from P_0 (the value of P at $\Phi=0$), and a cut-off mean absolute angle calculated with P cutoff at 4 times the cut-off mean. The corresponding values of ψ and T_ψ will be denoted by Φ_m, T_m; $\Phi_{\rm med}$, $T_{\rm med}$; $\Phi_{\frac{1}{2}}$, $T_{\frac{1}{2}}$; $\Phi_{1/e}$, $T_{1/e}$; Φ_0, T_0; $\Phi_{m;co}$, $T_{m;co}$.

We have in addition calculated the corresponding quantities for the distribution in the angle α between two adjacent chords, as a function of the mean number of scatterings in the track containing the two chords.[3] The quantities will be labeled α_m, C_m; $\alpha_{\rm med}$, $C_{\rm med}$; etc. (T for tangents and C for chords).

Another way of presenting multiple-scattering results is in terms of the "scattering constant." If some type of mean angle ψ is measured, the corresponding scattering constant $K_\psi=\sqrt{2}\psi p v/t^{\frac{1}{2}}$, where p and v refer to the scattered particle and t is the length of track. (The $\sqrt{2}$ is needed because, in the case of chord angles, our t includes *two* chords.) It is well known that K_ψ is a function only of $\frac{1}{2}t/\beta^2$. In terms of T_ψ (or C_ψ) we can write, using Eq. (4),

$$K_\psi^2=2p^2v^2\chi_\gamma^2\Omega T_\psi/t=(8\pi e^4Z'^2\Sigma_i N_i Z_i^2)T_\psi. \quad (6)$$

We find for the T's and C's:

$$T_m=\langle|\Phi|\rangle_m^2/\chi_\gamma^2\Omega=1.044+0.809\log_{10}\Omega; \quad (7a)$$

$$T_{\rm med}=|\Phi|_{\rm med}^2/\chi_\gamma^2\Omega=0.222+0.596\log_{10}\Omega; \quad (7b)$$

$$T_{\frac{1}{2}}=(\Phi_{\frac{1}{2}})^2/\chi_\gamma^2\Omega=0.035+1.831\log_{10}\Omega; \quad (7c)$$

$$T_{1/e}=(\Phi_{1/e})^2/\chi_\gamma^2\Omega=0.253+2.636\log_{10}\Omega; \quad (7d)$$

$$T_0=(\pi P_0^2\chi_\gamma^2\Omega)^{-1}=0.806+2.656\log_{10}\Omega; \quad (7e)$$

$$T_{m;co}=\langle|\Phi|\rangle_{m;co}^2/\chi_\gamma^2\Omega=0.418+0.818\log_{10}\Omega; \quad (7f)$$

$$C_m=\langle|\alpha|\rangle_m^2/\chi_\gamma^2\Omega=0.299+0.269\log_{10}\Omega; \quad (8a)$$

$$C_{\rm med}=|\alpha|_{\rm med}^2/\chi_\gamma^2\Omega=0.033+0.200\log_{10}\Omega; \quad (8b)$$

$$C_{\frac{1}{2}}=(\alpha_{\frac{1}{2}})^2/\chi_\gamma^2\Omega=-0.123+0.613\log_{10}\Omega; \quad (8c)$$

$$C_{1/e}=(\alpha_{1/e})^2/\chi_\gamma^2\Omega=-0.105+0.879\log_{10}\Omega; \quad (8d)$$

$$C_0=(\pi P_0^2\chi_\gamma^2\Omega)^{-1}=0.088+0.885\log_{10}\Omega; \quad (8e)$$

$$C_{m;co}=\langle|\alpha|\rangle_{m;co}^2/\chi_\gamma^2\Omega=0.090+0.272\log_{10}\Omega. \quad (8f)$$

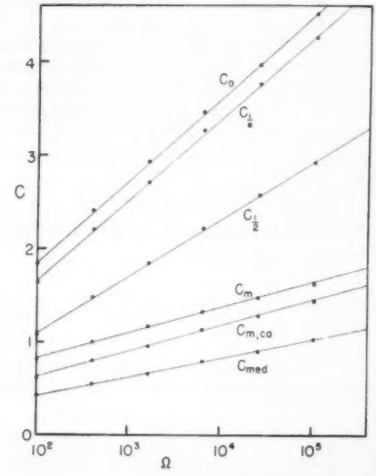


FIG. 2.

[7] S. T. Butler, Proc. Phys. Soc. (London) **63A**, 599 (1950).

Figs. 1 and 2 show the calculated results and fitted lines for the T's and C's.

With minor exceptions, the formulas agree with the calculations within 1 percent over the range of our tables, from $\Omega=100$ to $\Omega=84{,}000$ for the tangent distribution and 100 to 100,000 for the chord distribution. They were each fitted for six or seven values of Ω, by a least-square-relative-error method. The formula for T_{med} is off by 1.2 percent at one point, that for C_m by 1.3 percent at one point, and that for C_{med} by 1.6 percent at one point, with no other deviations as much as 1 percent.

The values for the mean absolute angle were compared with the analytic formula of Molière, and it was found that the linear relation is good to 1 percent out to $\Omega=10^7$, and to 3 percent to $\Omega=10^9$. The values for $T_{1/e}$, $T_{\frac{1}{2}}$, and T_0 are closely related to Molière's B, which obeys the linear relationship $B=1.153+2.583\log_{10}\Omega$ good to 3 percent out to $\Omega=10^9$. Hence we believe that all our formulas may be safely extrapolated for large Ω. Their accuracy is, however, not good for Ω much below 100.

Figure 3 shows a plot of B against log Ω, with Molière's values and some further calculations of ours, along with fitted line.

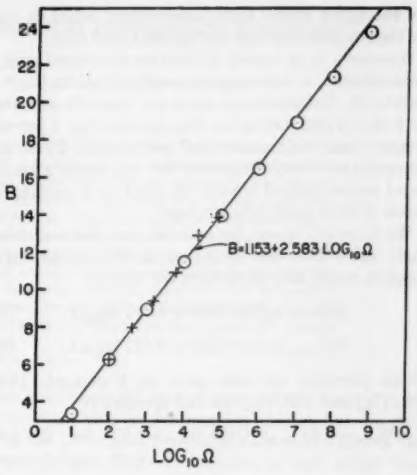


FIG. 3.

In the Gaussian approximation and in the limit of large Ω for the exact small-angle case, there is a relation between the C's and the T's. The factor as usually stated is $(3/2)^{\frac{1}{2}}$, but since Ω here for the chord case refers to two chords and the angles are squared, we find a factor 3. The coefficients of the logarithms in Eqs. (7) are all within 0.4 percent of 3 times the corresponding coefficients in (8). However, the constant terms show no such simple relationship.

We have not made calculations for the distribution of angles between "best-fit" or "smoothed-out" chords, as used by various investigators using photographic emulsions. However, Molière has dealt with this case in considerable detail.[8] The smoothing-factors for the Gaussian part of the calculation are given by Molière as $(3/2)^{\frac{1}{2}}$ for the chord case, $(35/26)^{\frac{1}{2}}$ for the angles between successive best-fitting (least-squares) lines, and $(70/61)^{\frac{1}{2}}$ for the angles between alternate best-fitted chords. We can thus approximate the effect of using successive best-fitted lines by multiplying our chord functions (8a)–(8f) by the factor $3/(35/13)=39/35$, and the effect of using alternate chords in half-cells by multiplying the tangent functions (7a)–(7f) (which give the nearest results) by 61/70. For this last calculation the length t in Eq. (5a) refers to the center-to-center distance of two alternate chords. Goldschmidt-Clermont[9] uses a smoothing factor of 0.96, for which $(61/70)^{\frac{1}{2}}=0.934$ is a closer approximation.

The extent to which the Gaussian approximation is valid can be shown by correspondences among the various T's. If we characterize a Gaussian distribution by its $1/e$ width $\Phi_{1/e}=w$, the various other quantities are given by $\langle|\Phi|\rangle_{\text{Av}}=w/\pi^{\frac{1}{2}}$; $|\Phi|_{\text{med}}=0.4769w$; $\Phi_{\frac{1}{2}}=w(\log_e 2)^{\frac{1}{2}}$; $(P_0\pi^{\frac{1}{2}})^{-1}=w$; and $\langle|\Phi|\rangle_{\text{Av, co}}=0.5614w$. Calculating w from the various T's by these formulas gives for the coefficient of $\log_{10}\Omega$, in order, (a) 2.593, (b) 2.637, (c) 2.642, (d) 2.636, (e) 2.656, (f) 2.595. For large Ω, then, the criteria (7b), (7c), (7d), and (7e) agree within 1 percent showing that the central part of the curve has indeed a Gaussian shape. The effect of the plural scattering tail is still felt to an extent of about 2 percent for large Ω.

It will be noted that we do not refer to a mean square angle. One of the simplifications used in the theory results in the fact that the single-scattering law does not possess a mean square unless a somewhat arbitrary cutoff is introduced. This fact has a negligible effect on the calculations unless one should wish a multiple-scattering mean square angle. However, the statistical errors introduced by using the mean square as an experimental criterion are considerable and make this calculation less useful than the simpler mean absolute angle.

On the other hand, a mean-square angle can be of use in calculating the errors involved in a series of repeated measurements. Specifically, if a series of chord angles are measured using $n+1$ cells on the same track, as in an emulsion, and a mean absolute angle calculated, the rms deviation of this result from the expected mean can be found by the well-known theorem on the variance of a sample mean for an arbitrary distribution—namely, that the rms deviation is $n^{-\frac{1}{2}}$ times the rms deviation expected for a single measurement. This result requires

[8] G. Molière, unpublished manuscript. I am indebted to Dr. Molière for communicating his results to me in prepublication form.

[9] Y. Goldschmidt-Clermont, Nuovo Cimento 7, 331 (1950).

the successive angles to be independent, which is true for the tangent case but not for the chord case.

However, in an earlier calculation concerned with a measurement of mean-square angle of scattering,[10] a correlated, Gaussian-approximation formula was used for 8 chords and yielded an rms spread about 5 percent greater than the uncorrelated calculation. Thus, the uncorrelated error calculation for the mean absolute chord angle method should be good to 5 percent or better if 10 or more cells are used.

We have calculated for both the tangent and chord angle cases the rms deviation of the cut-off mean absolute angle, and quote here the results

$$\langle\Phi^2\rangle_{\hbar;\,co}/\chi_\gamma^2\Omega=0.998+1.221\,\log_{10}\Omega; \qquad (7g)$$

$$\langle\alpha^2\rangle_{\hbar;\,co}/\chi_\gamma^2\Omega=0.202+0.422\,\log_{10}\Omega. \qquad (8g)$$

These formulas are also good to 1 percent. Using (7f), (7g) and (8f), (8g) we find specifically

$$[\langle\Phi^2\rangle_{\hbar;\,co}-\langle|\Phi|\rangle_{\hbar;\,co}^2]^{\frac{1}{2}}/\langle|\Phi|\rangle_{\hbar;\,co}=0.82 \quad \text{for} \quad \Omega=10^2$$
$$=0.77 \quad \text{for} \quad \Omega=10^5;$$

$$[\langle\alpha^2\rangle_{\hbar;\,co}-\langle|\alpha|\rangle_{\hbar;\,co}^2]^{\frac{1}{2}}/\langle|\alpha|\rangle_{\hbar;\,co}=0.80 \quad \text{for} \quad \Omega=10^2$$
$$=0.77 \quad \text{for} \quad \Omega=10^5.$$

Thus, no serious misestimation of error will be made if we use for n angular measurements of any sort on a track, a standard deviation for the arithmetic mean of $0.85n^{-\frac{1}{2}}$ times the mean.

For the case of Ilford G-5 emulsions, we have calculated the scattering constant K. In agreement with Snyder,[11] we find that $\Sigma_i N_i Z_i^2=3.70\times10^{25}$, or $\langle Z^2\rangle_{\hbar}^{\frac{1}{2}}=21.34$. However, using Eq. (4) instead of an average of $Z^{4/3}$ we find, combining with Eq. (5) and $\beta=1$, that $\Omega=631s_0$ where s_0 is the cell-length $t/2$ in units of 100 microns. This result is equivalent to using $Z=40.6$ in Eq. (3b), and differs from the value $z=1248s_0$ of reference 11, which was calculated from Eq. (2a) without the factors 1.13^2 and $1.13+3.76\gamma^2$, and involved $\langle Z^{4/3}\rangle_{\hbar}=(16.22)^{4/3}$. For $s_0=1$, $\Omega=631$, we find instead of the value[11] 26.8: (a) complete mean value, $K_m=25.8$ from (8a); (b) cut-off mean value, $K_{mco}=23.2$ from (8f). If we wish to compare the results of using Eqs. (7) and a factor $\frac{1}{3}$, we find: (c) complete mean value from $T_m/3$, $K_m'=26.4$; (d) cut-off mean value from $T_{mco}/3$, $K_{mco}'=23.9$.

With somewhat less reliability, we can calculate scattering constants using Molière's smoothing factors. For best-fitted lines (assumed to be the same as a least-squares fitting would yield), we have: (e) complete mean value from $(39/35)C_m$, $K_{m-\text{smooth}}=27.2$; (f) cut-off mean value from $(39/35)C_{mco}$, $K_{mco-\text{smooth}}=24.5$.

Finally, for angles between alternate smoothed chords at separations $2s$, we have, dividing by 2 to reduce to cell-length s: (g) complete mean value from $(61/140)T_m$, $K_{m-\text{alt}}=30.2$; (h) cut-off mean value from $(61/140)T_{mco}$, $K_{mco-\text{alt}}=27.4$. The use of $Z_i^2+Z_i$ instead of Z_i in these results would increase $\langle Z^2\rangle_{\hbar}$ by 0.29 percent and would have a negligible effect on Ω. Hence all the results above can be increased by about 0.1 because of this effect. The variation of K with cell-length is to be found, of course, from the appropriate one of Eqs. (7) or (8).

The variation of Ω with β^2 is given to sufficient accuracy by Fig. 2 of the paper by Gottstein *et al.*[12] The values given by this figure must be multiplied by 2 for use in our formulas.

We have compared our results with those of Williams' scattering theory[13] as discussed by Voyvodic and Pickup.[14] The results of this theory can be written in similar form to Eqs. (7). In particular, the mean and cut-off mean values of Williams' theory fit the following formulas:

$$T_m{}^W=1.63+0.79\,\log_{10}\Omega, \quad T_{mco}{}^W=1.18+0.79\,\log_{10}\Omega.$$

The differences from the present results will evidently be greatest for small Ω.

Work is progressing on formulas similar to Eqs. (7) for the spatial (unprojected) tangent angle distribution, and will be reported later.

We should like to take the opportunity here to point out that the graphs of the two Snyder-Scott distributions printed in reference 3, Figs. 5 and 6, have had their legends interchanged.

Note added in proof: Recent, unpublished words by S. Olbert on the inclusion in Molière's theory of the effect of the finite size of nuclei indicates that the tails of the multiple scattering distribution are considerably reduced by this effect. Thus our results may need modification, especially for the mean absolute angles; to a lesser extent for the cut-off mean median and $\langle P_0w\rangle^{-\frac{1}{2}}$ values; and probably very little for the $\frac{1}{2}$ and $1/e$ widths. I am indebted to Dr. Bruno Rossi for communicating these results to me.

ACKNOWLEDGMENTS

Credit is hereby gratefully given to Miss Jean Snover, Miss Theresa Danielson, Miss Rika Sarfaty, and Miss Gertrude Nissenbaum for the numerical calculations.

[10] W. T. Scott, Phys. Rev. **75**, 1763 (1949).
[11] H. S. Snyder, Phys. Rev. **83**, 1068 (1951).
[12] Gottstein, Menon, Mulvey, O'Ceallaigh, and Rochat, Phil. Mag. **42**, 716 (1951).
[13] E. J. Williams, Phys. Rev. **58**, 292 (1940).
[14] E. Pickup and L. Voyvodic, Phys. Rev. **84**, 1190 (1951).

Fast Protons from the Capture of π^- Mesons in Photographic Emulsions*†

FRANK L. ADELMAN
Radiation Laboratory, Department of Physics, University of California, Berkeley, California
(Received October 1, 1951)

The spectrum of fast protons from 1631 π^- meson absorptions in Ilford G-5 and C-2 emulsions has been studied. It has been found that (16.1±1.8) percent of all π^- meson absorptions are accompanied by protons of energy greater than 20 Mev, and that (9.5±1.3) percent of all absorptions are accompanied by protons of energy greater than 30 Mev. The proton energies were estimated by grain counting in comparison with π^- mesons of known residual range. Allowance was made for the variation of grain density with position and depth in the emulsion. The energy spectrum is fairly flat from 30–70 Mev, but only three events were found above 70 Mev. Examination of stars from which the fast protons were emitted shows more prongs ejected opposite to the direction of the fast proton than are accounted for by evaporation from a recoiling nucleus. The energy spectrum favors the capture of the meson by a proton, with the recoil momentum taken by a small, variable number of nucleons. Furthermore, the evidence suggests that the low energy star prongs result from knock-on collisions of the primary nucleons with other nucleons inside the nucleus.

I. INTRODUCTION

THERE have been several attempts recently to discover the mechanism by which a π^- meson is captured by a complex nucleus. In the earliest models, the whole nucleus participated in the capture, and the entire rest of the mass of the π^- meson (about 140 Mev) was available for excitation of the nucleus. The evaporation models with an excitation energy of 140 Mev[1,2] fail to predict the observed energy spectrum[3,4] of the emitted charged particles. Furthermore, the observed average excitation energy of the meson stars was about 100 Mev,[3–5] considerably less than the total available energy. In 1949 Marshak[6] proposed a model in which the π^- meson interacts with a single proton of the nucleus, and the recoil momentum is taken up by a neighboring nucleon. The preliminary calculations indicated that more than 10 percent of the π^- absorptions should be accompanied by the emission of a proton of energy greater than 30 Mev. More detailed calculations by Fujimoto *et al.*[7] and by Tamor[8] supported this prediction. In order to check this prediction, the author,[9] Cheston and Goldfarb,[10] and Menon *et al.*[4] examined π^- meson endings in photographic emulsions capable of recording tracks of charged particles even at minimum ionization. They observed that fast protons[11] were indeed emitted from meson-induced stars, a result inconsistent with an evaporation process, but the data were too inconclusive to confirm or reject any of the specific hypotheses. The author's first study was essentially exploratory, with very few statistics and a crude method of energy estimation. Menon *et al.* and Cheston and Goldfarb had better statistics and more precise methods of energy determination. However, in each case the criterion of energy, the grain density, varied relatively slowly with proton energy in the neighborhood of 30 Mev. Furthermore, since the variation of observed grain density with position in the plate[12] and with depth in the emulsion was not taken into account, the number of protons of energy greater than 30 Mev could not be reliably determined.

II. EXPERIMENTAL ARRANGEMENT

The photographic plates were exposed in the 184-inch Berkeley cyclotron as shown in Fig. 1. They were put a relatively large distance from the target in order to reduce the background due to neutral particle from the target. At this distance most of the mesons which could reach the plates had enough energy (~30 Mev or more) to pass completely through the pack of plates. Therefore, a ¾-inch aluminum absorber was placed in front of the plates to slow down the mesons.

Ilford G-5 emulsions were chosen for this experiment

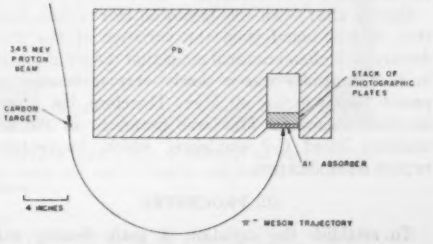


FIG. 1. Schematic diagram of apparatus for exposing plates.

* Submitted in partial fulfillment of the degree of Doctor of Philosophy in the Department of Physics, University of California, Berkeley, California.

† This work was performed under the auspices of the AEC.

[1] E. Clementel and G. Puppi, Nuovo cimento **6**, 494 (1949).

[2] Y. Fujimoto and Y. Yamaguchi, Prog. Theor. Phys. **4**, 468 (1949).

[3] D. Perkins, Phil. Mag. **40**, 601 (1949).

[4] Menon, Muirhead, and Rochat, Phil. Mag. **41**, 583 (1950).

[5] J. Heidmann and L. LePrince-Ringuet, Compt. rend. **226**, 1716 (1948).

[6] R. Marshak, Echo Lake Symposium on Cosmic Radiation (1949).

[7] Fujimoto, Hayakawa, and Yamaguchi, Prog. Theor. Phys. **4**, 576 (1949).

[8] S. Tamor, Phys. Rev. **77**, 412 (1950).

[9] F. Adelman, Phys. Rev. **78**, 86(A) (1950).

[10] W. Cheston and L. Goldfarb, Phys. Rev. **78**, 683 (1950).

[11] In the rest of this paper, the term "fast proton" will refer to a proton of energy greater than 30 Mev.

[12] K. Bowker, Phys. Rev. **78**, 87(A) (1950).

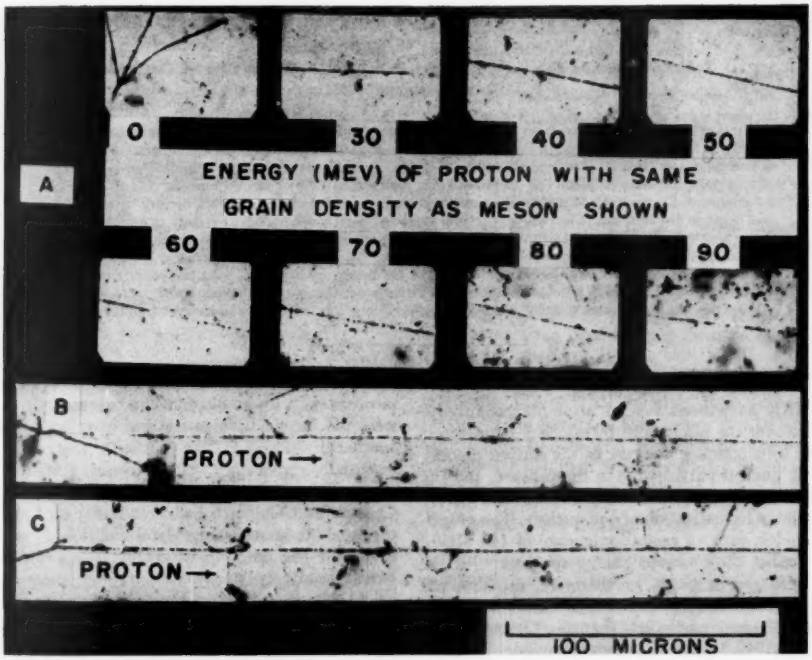


FIG. 2. (A) Photomicrograph of meson track taken at various velocities in G-5 plate. A proton at the same velocities would have the grain densities and energies shown. (B) 4-prong meson star, one of whose prongs is a 58-Mev proton. (C) Meson star with single prong, a 55-Mev proton. An electron is also associated with this event.

so that the most energetic protons might be observed. Since these plates had acquired a heavy background of low energy electrons by the time they had reached Berkeley, it was desirable to eradicate the latent images by accelerated fading.[12] By storing the plates in an oven at about 33°C over a pan of water for about 50 hours and then drying, it has been found possible to eliminate 90 percent of the latent image electrons without any noticeable loss of sensitivity. The plates were exposed and developed as soon as possible after the background eradication.

Shortly after work was begun on this project, however, it was found from the variation of the grain density with meson residual range that the grain density in G-5 emulsions was a slowly varying function of proton energy below 40 Mev. Therefore, the proton spectrum below 60 Mev was examined in the less sensitive Ilford C-2 emulsions, which, incidentally, require no eradication.

III. PROCEDURE

To establish the variation of grain density with energy, mesons were grain-counted up to a distance of 3.7 mm from the end of their ranges, at which point the grain density was the same as that of a 90-Mev proton. It was found that the differential grain count could be treated as the product of two factors, one depending only upon energy and the other, only upon the depth in the emulsion.

The differential grain count was found to vary as $1/(E_{\text{proton}})^{\frac{1}{2}}$ in the G-5 plates in the range 45–90 Mev and in the C-2 plate in the range 20–60 Mev. Protons of energy greater than 60 Mev in the C-2 plates were not included, as the fluctuations of grain count due to the density of random grains are too large. Corrections were applied for the variation of grain count with depth in the emulsion.

IV. RESULTS

A total of 1631 meson absorptions were examined, 992 in G-5 emulsion and 639 in C-2 emulsion. Typical examples and grain densities of protons[14] of energy

[12] J. Spence, private communication.

[14] The grain density of a 20-Mev proton is equal to that of a 40-Mev deuteron or of a 60-Mev triton. Therefore the term "proton" includes deuterons of twice the energy and tritons of three times the energy in question. There are a few cases where positive identification can be made from the range of the particles, as will be discussed below.

greater than 20 Mev are given in Figs. 2 and 3. The energy spectrum of these protons, after the application of the geometrical corrections, is shown in Fig. 4; the standard deviation in the energy of a proton varies from 2½ Mev at 20 Mev to about 6 Mev at 70 Mev. (16.1±1.8) percent of all meson absorptions are accompanied by protons of energy greater than 20 Mev, and (9.5±1.3) percent of all absorptions are accompanied by protons of energy greater than 30 Mev. In terms of ejected particles, (10.3±1.2) percent of all star prongs are protons of energy greater than 20 Mev, and (6.0 ±0.8) percent are protons of energy greater than 30 Mev. These numbers are in agreement with those of other workers.[4,10]

A precise prong spectrum[15] for meson induced stars may also be given by use of the data of this study. Since Adelman and Jones,[16] and Menon *et al.*[4] used the same conventions[17] for star prongs as the author, their data may be combined with that found here to give the prong spectrum of Table I, based upon 3366 meson stars.

V. DISCUSSION

When these results are compared with the theoretical predictions of π^- meson capture,[4,6,7,8,18] it may be seen that most of the models yield higher energy protons or larger numbers of protons (or both) than have been found here.

The model of Menon *et al.*,[4] however, does not seem inconsistent with the present data. They assumed that the primary interaction takes place among 3 or 4 nucleons, and that the energy of those nucleons which do not escape is used in heating the nucleus. For their preliminary calculations, they assumed an α-particle model, for which calculations had been made by Ruddlesden and Clark.[19] The predictions of this preliminary computation are an excitation energy of 120 Mev, which is somewhat high, and a peak in the number of protons between 30 and 40 Mev, for which there is no evidence in Fig. 4. However, when the interactions of the smaller numbers of nucleons are taken into account, the mean excitation energy should decrease and the peak should be smoothed out. Thus,

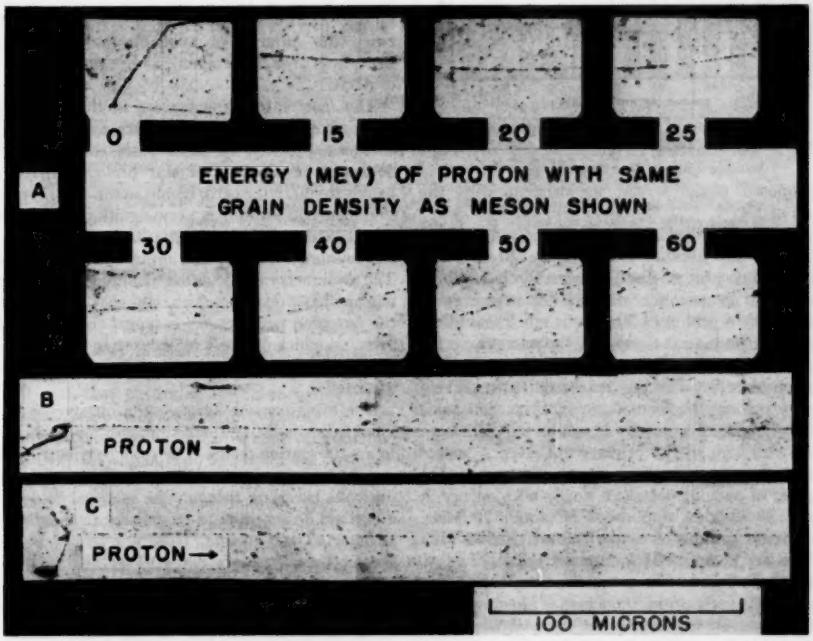


FIG. 3. (A) Photomicrograph of meson track taken at various velocities in a C-2 plate. A proton at the same velocities would have the grain densities and energies shown. (B) 2-prong meson star, one of whose prongs is a 20¼-Mev proton. (C) 3-prong meson star, one of whose prongs is a 56½-Mev proton.

[15] I.e., the frequency distribution of stars of a given number of prongs.
[16] F. Adelman and S. Jones, Science **111**, 226 (1950).
[17] Any group of grains (except an electron track) leaving the terminus of a meson with a well-defined direction is a star prong.
[18] Fujimoto, Takayanagi, and Yamaguchi, Prog. Theor. Phys. **5**, 498 (1950).
[19] S. Ruddlesden and A. Clark, Nature **164**, 487 (1949).

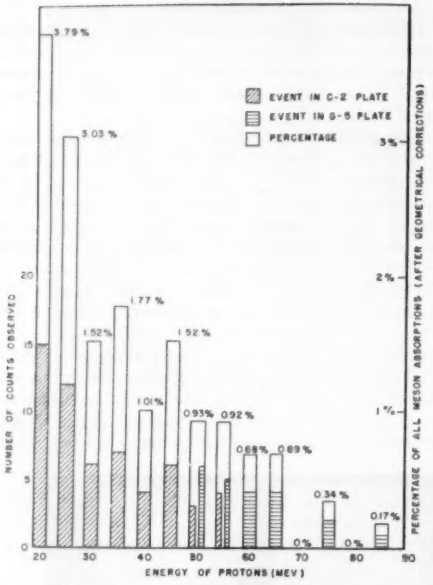


FIG. 4. Energy spectrum of emitted protons of energy greater than 20 Mev, before and after geometrical corrections.

their proposed model is not inconsistent with the spectrum of Fig. 4.

In Fig. 4 the number of protons in the range 20–30 Mev is twice the number in the range 30–40 Mev. Since a 20-Mev proton could reasonably have come from a nuclear evaporation with excitation energy of 100 Mev, while a proton of 30 Mev is much less likely to have come from such a process, the energy spectra of the protons from the primary and secondary processes overlap considerably. This implies either that relatively low energy primary nucleons occur often or that some of the lower energy protons are knock-on particles from collisions with high energy primary nucleons.

Another characteristic of the energy spectrum is that the number of protons decreases slowly with energy in the range 30–70 Mev, and drops off above 70 Mev; there are fewer protons of energy greater than 70 Mev than there are in any 5-Mev interval below 70 Mev. Thus relatively few of the π^- absorptions take place with the creation of a primary nucleon of energy greater than 70 Mev, while nucleons occur often with energies lower than this figure. Therefore, the primary act must take place often among two or more nucleons, with a deuteron or triton seldom acting alone as the recoil particle. On the other hand, three of the emitted particles ended in the emulsion. Two of these were actually protons by range and grain density, and the third was a 54-Mev deuteron. In addition, one particle was found which was probably a triton of 30 Mev, and another was either a 23-Mev proton with abnormally high grain density or a 31-Mev deuteron with abnormally low grain density. Neither of these is included in the data. Thus it appears that deuterons and tritons actually take part in the primary process; if they were knock-on particles, the energetic primaries required would lead to a larger number of protons of energy greater than 50 Mev than in the range 30–50 Mev,[20] in contradiction to Fig. 4. This conclusion does not contradict the earlier conclusion of this paragraph, since there is not enough data to observe the relative numbers of protons, deuterons, and tritons.

The prong spectrum of the stars in which a fast proton appears is quite similar to the normal prong spectrum of π^- mesons absorbed in photographic emulsions (Table I). If one takes into account the limitations of the comparison of prong spectra with poor statistics and the lack of detailed knowledge of the process, one may still state reasonably that most of the π^- absorptions which are not accompanied by protons of energy greater than 20 Mev are accompanied by neutrons of similar energies. This would seem to indicate that the number of neutrons available as the recoil particle is twice the number of available protons (since one neutron is always involved), as in the α-particle model of nuclei.

The most interesting feature of the data appeared upon an examination of the stars associated with the energetic protons. In these stars the observed ratio between the numbers of star prongs emitted in the backward and forward hemispheres (with respect to the energetic proton) is very significantly higher than that calculated with generous assumptions, ascribing the asymmetry to the motion of the recoiling nucleus. The asymmetry is calculated assuming that the residual nucleus has a mass twelve times that of a proton, and the outgoing particles are assumed to have the lowest velocity which is used in measuring the asymmetry. All particles are assumed to be either protons or α-particles.

The disagreement between the observed and calculated ratios is, in fact, striking, and leads immediately to an alternative model. The secondary star prongs are assumed to be due mainly to impacts with primary nucleons traveling through the nucleus. Seldom would a particle be expected to be emitted in the direction of an emitted fast proton. However, one can no longer assume that there are two 70-Mev nucleons traveling in opposite directions, for then the observed energy spectrum would have more high energy protons.[20] ‡

[20] M. Goldberger, Phys. Rev. **74**, 1269 (1948).

‡ *Note added in proof:* It has been pointed out by Dr. H. York that the calculations of Goldberger (see reference 20) are not consistent with experiment, so that the conclusions based upon this computation are weakened considerably. (See J. Hadley and H. York, Phys. Rev. **80**, 345 (1950). (The experimental data are, unfortunately, not strictly applicable to the problem of meson absorption.) Nevertheless, since the energy spectrum of the primary nucleons in the two-nucleon model will extend well above 70 Mev when the initial momentum distribution is taken into account, it is puzzling that so few protons of energy greater than 70 Mev are seen.

Therefore it is proposed that the primary interaction takes place among a small, but variable, number of nucleons (most often 2, 3, or 4), as in Menon *et al.*[4] This permits the emission of protons up to 70 Mev with reasonable probability, and protons up to 95 Mev occasionally. These nucleons have a spread of energies which favors the emission of low energy knock-on particles. Since the mean free path of such a nucleon is of the order of the nuclear radius for an element like Ag,[20] each nucleon will collide, on the average, one time inside the nucleus. But most of the knock-on particles will not be able to escape from the nucleus. Qualitatively, therefore, about one extra prong is expected, on the average, to accompany an energetic prong. This figure is consistent with the observed number of star prongs.

A test of this hypothesis would be to calculate the energy and angular distribution of the knock-on particles. The Monte Carlo method of calculation appears to be most suitable, but it will not be attempted in this paper.

TABLE I. Prong spectrum of π^- meson stars.

Number of prongs	1	2	3	4	5	6
Percent of stars	32.4±1.0	32.7±1.0	22.3±0.8	10.5±0.6	2.0±0.2	0.1±0.1

VI. ACKNOWLEDGMENTS

I should like to thank Dr. W. Barkas, Dr. H. Bradner, Dr. C. Richman, and Dr. R. L. Thornton for their guidance and encouragement, and to express my gratitude to the other members of the Film Program at the Radiation Laboratory for their cooperation and for many fruitful discussions. I am also indebted to Mr. A. Oliver for the preparation of Figs. 2 and 3, and to Miss Betty Summers for the other figures.

PHYSICAL REVIEW VOLUME 85, NUMBER 2 JANUARY 15, 1952

Limits of Stability of the Heavy Elements and the Effects of Nuclear Shell Structure*

P. S. GREIFINGER
Cornell University, Ithaca, New York

(Received September 14, 1951)

The systematics of beta-decay, alpha-emission, and spontaneous fission will be investigated in order: first, to determine the theoretical limits to the periodic table, and secondly, to examine the effects of nuclear shell structure on the end of the periodic table. Curves of beta-stability and lifetimes against alpha-emission and spontaneous fission will be drawn for the heaviest elements and will be extrapolated in order to yield information about the stability or form of instability of the transuranic elements. Nuclear shell structure will be seen to be responsible for the stability of Th^{232} and U^{238} as well as the instability of the group of highly radioactive elements beginning with polonium.

I. INTRODUCTION

URANIUM is the heaviest known naturally occurring element on a macroscopic scale. Presumably, when the earth was formed ($\sim 3\times 10^9$ years ago), several isotopes, now unknown, were present. The only nuclei still occurring naturally are those whose half-lives are appreciable compared to 10^{8-9} years.

The beta-stable isotopes occur in a relatively narrow band; those nuclei lying outside this band are unstable with respect to positron emission or K-capture on the one side, and negative beta-emission on the other. This band would theoretically continue without limit onward towards higher and higher atomic number. However, the beta-stable transuranic elements do not occur in nature because of instability against alpha-decay or fission.

The manner in which an unstable nucleus is most likely to have decayed is governed by the lifetimes against the three forms of decomposition. At present there is no isotope known which has had fission as its main mode of decay.

Existing data will here be used to draw curves of beta-stability and lifetimes against alpha-decay and fission. These curves will yield information on the effects of nuclear shell structure on the end of the periodic table. To obtain information about the stability or form of instability of the transuranic elements, these curves will be extrapolated. In performing the extrapolations, it is assumed that nuclear shell structure alone is responsible for all major irregularities in the nuclear masses.

II. BETA-STABILITY

As a result of the dependence on the even and odd character of the proton and neutron numbers, all nuclides may be divided into three general classes, those of odd A, those of even A and even Z, and those of even A and odd Z. For each class, the nuclear mass as a function of A and Z lies on a surface having the general shape of a valley. A curve connecting the values of Z_A, the coordinate of the bottom of the "odd A" valley, as a function of A is often called the line of beta-stability. All beta-stable odd-A nuclei lie within $\frac{1}{2}Z$ units of this curve. The region of stability of the even-even nuclides is bounded by limits approximately symmetrical with respect to Z_A. The curves representing Z_A and the region of beta-stability are here indicated on

* Part of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Physics in the Graduate School of Cornell University.

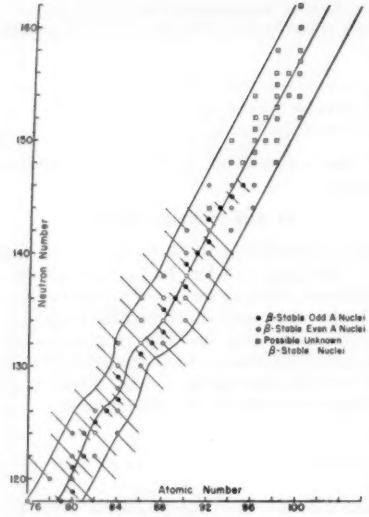


FIG. 1. The limits of beta-stability for the heavy nuclei are plotted on an N–Z plane.

an N–Z plane (Fig. 1), where 45-degree lines represent lines of constant A.

The method used to draw the curves, first suggested by Bohr and Wheeler[1] and later elaborated upon by Kohman,[2] is here carried out for $A>194$. Since the beta-stability of the short-lived alpha-emitters cannot be determined in all cases, the curves are somewhat rougher approximations in this region than they would be for lower A. However, one can get a good idea of the general trends of the curves, and even a sharp indication of the irregularity in the region of 126 neutrons and 82 protons.

In order to make predictions concerning the beta-stability of nuclei with $A\gtrsim240$, the curves must be extrapolated in some fashion. For $A\gtrsim228$, if possible slight local irregularities are ignored, the curves can be represented by straight lines of the same slope. This slope is slightly greater than the average slope up to this point, because of the tendency towards an increasing N–Z ratio with increasing number of particles due to the Coulomb repulsion of the protons. These lines are continued onward toward higher A with a further slight increase in slope. Any irregularities which would occur in the shape of the curves cannot be foreseen; however, because another completed shell will not occur for a long way off, the irregularities should be of such a nature as not to affect the choice of likely beta-stable nuclides. Possible beta-stable nuclei for $A>242$ are indicated in Fig. 1 as squares and are chosen to stay within the limits of the extrapolated curves. Odd-odd nuclei are excluded, and except where there is considerable doubt as to which of two to pick, only one nucleus for each odd A is included.

[1] N. Bohr and J. A. Wheeler, Phys. Rev. **56**, 426 (1939).
[2] T. P. Kohman, Phys. Rev. **73**, 16 (1948).

III. ALPHA-STABILITY

Most nuclei with $A\gtrsim140$ are energetically capable of emitting alpha-particles, but are stabilized by the nuclear potential barrier resulting from the high charge. As the decay energy of the alpha-particle becomes small, the lifetime of the alpha-radioactive nucleus becomes extremely large. There is no observed alpha-activity where the kinetic energy of the alpha-particle is $\lesssim2$ Mev, since the maximum half-life which can be detected by present methods is $\sim10^{21}$ years.

In order to make predictions concerning the alpha-stability of unknown isotopes of the transuranic elements, one can use existing data on alpha-energies and lifetimes, and examine the trends to make reasonable extrapolations.

There are several ways in which to correlate alpha-decay energies, some of which have already been used.[3] Alpha-energy contours are here drawn on an N–Z plane utilizing data from almost 100 alpha-decaying species.[3–5] Contours corresponding to 4, 5, 6, 7, 8, and 9-Mev alpha-particle energies are drawn (Fig. 2). The contours appear to be essentially parallel straight lines until $N=122$ and $Z=82$, where a very dramatic break begins to occur. Abnormally high alpha-energies are obtained for nuclei with $128\leq N\leq134$ and $Z=84$. The configuration of 126 neutrons and 82 protons is revealed as an abnormally stable structure compared with higher neutron and proton numbers.

For $N\gtrsim140$, the contours again appear as parallel straight lines, but are displaced from the position they had before the break. In keeping with the assumption that completed shells are responsible for all major irregularities and that another completed shell will not occur for a long way off, the contours are extrapolated toward higher A by continuing the straight lines onward with unchanged slope.

The energy contours obtained with the use of the Weizäcker-Fermi semi-empirical mass formula[6] are also drawn in Fig. 2. These contours are a series of parallel straight lines, parallel to the experimental curves until the closed-shell interruption, after which ($N\gtrsim140$) the two sets of curves again resume an almost parallel course. Thus, except for the sudden shell structure irregularities, the mass formula predicts the general behavior of the alpha-energy contours in the region where some data exists, and acts as a further justification for the extrapolation performed where there is no data.

[3] Perlman, Ghiorso, and Seaborg, Phys. Rev. **77**, 26 (1950).
[4] G. T. Seaborg and I. Perlman, Revs. Modern Phys. **20**, 585 (1948).
[5] W. H. Sullivan, *Trilinear Chart of Nuclear Species* (John Wiley and Sons, Inc., New York, 1949).
[6] M. O. Stern, Revs. Modern Phys. **21**, 316 (1949).

In order to predict the stability against alpha-decay of a given nucleus one must be able to determine the order of magnitude of its lifetime against alpha-emission. Thus, what is needed is a series of alpha-lifetime contours which can be extrapolated towards higher A.

If one plots the log of the half-life against alpha-decay *versus* the energy of the emitted alpha-particle, one obtains a family of curves, the parameter Z being held constant for each curve. Unfortunately, these curves are defined only for the even-even nuclei, all other nuclear species falling off the curves, and showing, in general, abnormally long lifetimes.[3]

Using these curves and the energy contours of Fig. 2, lifetime contours for the even-even nuclei are drawn on an N–Z plane in Fig. 3. Contours corresponding to half-lives of 10^{-1}, 10^{2}, 10^{5}, 10^{7}, and 10^{9} years are drawn for $Z \geq 92$.

IV. SPONTANEOUS FISSION

The phenomenon of fission is much more difficult to deal with than either alpha- or beta-decay. There exists little experimental data and no theory which fits that experimental information which is known. To some extent, the phenomenon of fission permits a classical interpretation in terms of the liquid drop model of the nucleus.[1,7] The theory is inadequate in many respects. Fission is, in general, not into symmetric fragments, and the fission thresholds do not vary smoothly. It

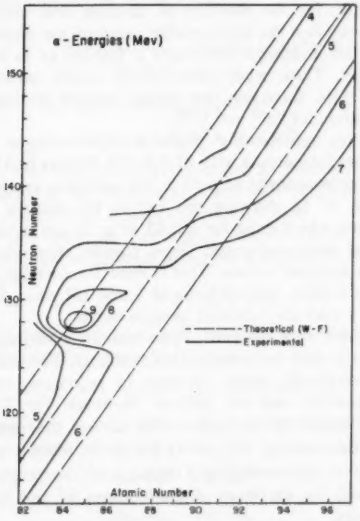


FIG. 2. Contours representing the total energies available to emitted alpha-particles are plotted on an N–Z plane. The heavy lines are based on experimental energies, while the dotted line contours are obtained by computing alpha-energies from the Weizsäcker-Fermi mass formula.

[7] S. Frankel and N. Metropolis, Phys. Rev. **72**, 914 (1947).

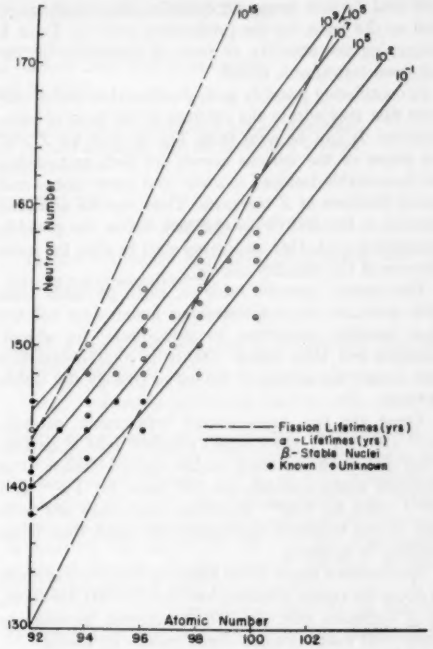


FIG. 3. Known and possible unknown beta-stable nuclei and contours of lifetimes against alpha-emission (defined only for the even-even nuclei) and spontaneous fission for the transuranic nuclei.

seems that the best one can do at present in order to make any predictions concerning the order of magnitude of fission lifetimes is to draw upon the general aspects of the theory in order to make qualitative predictions, and fit the numbers as well as one can to experimental lifetimes.

According to the theory, except for slight variations with A, contour lines of constant Z^2/A should be constant fission lifetime contours.

An estimated lower limit of the half-life of U^{235} ($Z^2/A = 36$) for spontaneous fission is about 10^{15} years.[8] The spontaneous fission half-life of Cm^{242} ($Z^2/A = 38.1$) is about 10^{5} years.[9] Contour lines are drawn (Fig. 3) for $Z^2/A = 36$ and $Z^2/A = 38.1$ with the assigned lifetimes of 10^{15} and 10^{5} years, respectively.

V. CONCLUSION

In Fig. 3, the known and possible unknown beta-stable nuclei are plotted as closed and open circles, and the alpha- and fission lifetimes are represented by

[8] W. E. Stephens, *Nuclear Fission and Atomic Energy* (The Science Press, Lancaster, Pennsylvania, 1948), p. 12.

[9] Hanna, Harvey, Moss, and Tunnicliffe, Phys. Rev. **81**, 466 (1951).

solid and broken lines, respectively. This diagram is used as the basis for the predictions made in Table I concerning the stability or form of instability of the unknown transuranic nuclei.

In considering possibly geologically stable nuclei, one must first stay within the confines of the band of beta-stability. It can be seen from Fig. 3, that for $Z>92$ the slopes of the various curves are such as to bring the beta-stable isotopes to lower and lower alpha- and fission lifetimes as Z increases. Thus, the list of stable isotopes is brought to an end long before the possible occurrence of another completed shell to alter the main features of the stability curves.

The natural periodic table on earth probably ends with uranium. Since spontaneous fission does not act as a possible competitor to alpha-decay for atomic numbers less than about 100, it is alpha-instability that causes the ending of the list of geologically stable isotopes.

There are two undiscovered transuranic isotopes, Pu^{244} and Cm^{250}, which might possibly exist in nature. They are both certainly stable against fission. The probable alpha-lifetimes are 10^9 years for Pu^{244}, and 10^{7-8} years for Cm^{250}. However, since they are both just on the border of the beta-stable band, their beta-stability is in doubt.

Spontaneous fission is not likely as the principal mode of decay for atomic numbers less than $Z=100$. However, as Z increases past $Z=100$, it becomes increasingly likely that many isotopes decay mainly by fission.

The alpha-stability of the heavy elements is very strongly affected by the phenomenon of shell structure, which results in completed shells with 82 or 126 particles. It is this phenomenon which is responsible, first, for the sudden appearance of a group of highly radioactive elements beginning with polonium ($Z=84$) and

TABLE I. The unknown beta-stable transuranic isotopes, their probable main mode of decay and log of estimated lifetimes (years) against this form of decay.

Z	A	Main decay form	Log lifetime
94	242	alpha	7
	244	beta, alpha	?, 9
95	243	alpha	4
96	244	alpha	1
	245	alpha	2
	246	alpha	3.5
	247	alpha	5
	248	alpha	6
	250	beta, alpha	?, 7.5
97	247	alpha	1
	249	alpha	3
98	246, 248	alpha	−3 to 6
	250–254		
	256		
99	253	alpha	0
	255	alpha	2
100	252, 254, 256	fission, alpha	−4 to 5
	257, 258, 260		
	262		

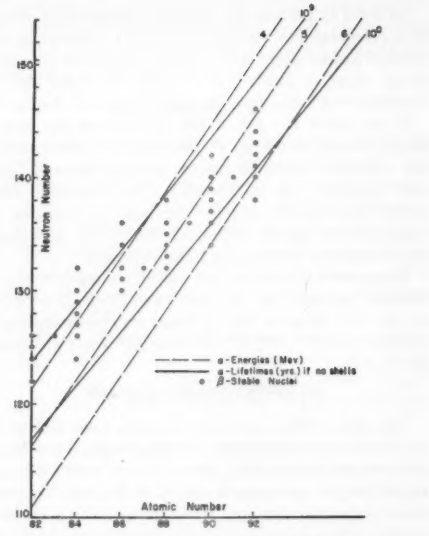


FIG. 4. Contours of energies of emitted alpha-particles and lifetimes against alpha-emission for the elements at the high end of the periodic table as they would be if there were no shell structure.

secondly, for the stability of thorium and uranium isotopes, since the alpha-energy contours are displaced after the interruption in such a fashion as to favor stability. Thus, while there are no stable nuclei for $84\leq Z\leq 89$, there are two stable isotopes of thorium and uranium (Th^{232} and U^{238}).

Curves representing probable alpha-energies and lifetimes for isotopes with $82\leq Z\leq 92$, if there had been no completed shells with 82 or 126 particles, are drawn in Fig. 4. In drawing the curves the straight line contours which exist for $Z\lesssim 83$ (Fig. 2) are extended onward with unchanged slope, parallel, therefore, to the Weizsäcker curves. This is done for energies of 4, 5, and 6 Mev, and lifetimes of 1 and 10^9 years. Beta-stable nuclei are indicated as open circles.

If there were no shells, alpha-activity would appear gradually with increasing atomic number. There would be geologically stable isotopes of polonium, radon, and possibly one for radium. However, for $Z>88$, there would be no such stable nuclei, the periodic table thus ending with either Rn or Ra instead of U. Moreover, the possibility of fission would be completely obscured by the short alpha-lifetimes of the heavy elements.

The author wishes to thank Professor Philip Morrison of Cornell University, who suggested the problem, for his continued assistance and encouragement.

The Force between Particles in a Nonlinear Field Theory

NATHAN ROSEN AND HERBERT B. ROSENSTOCK*
Department of Physics, University of North Carolina, Chapel Hill, North Carolina
(Received August 20, 1951)

It is shown for a general class of scalar nonlinear classical field theories, in which singularities are excluded and particles are represented by small regions in which the field is intense, that the interaction between two particles is described by the Yukawa potential at large distances.

1. INTRODUCTION

IN recent years a number of attempts have been made to set up a nonlinear field theory of elementary particles in which singularities are excluded.[1-4] In such a theory a particle is represented by a small portion of space in which some function representing matter density has a large value. Such a theory can be expected to be free from at least some of the divergences present in theories in which particles are represented by points, i.e., by singularities. It also has the advantage that the equations of motion of a particle are a consequence of the field equations.

It is of interest to know something concerning the nature of the force which one particle exerts on another according to a nonlinear field theory. For this purpose a very simple classical nonlinear theory has been investigated. The field in this case is taken to be described by a complex scalar ψ, the behavior of which is determined by the Lagrangian density function

$$\mathfrak{L}=-\frac{\partial\psi}{\partial x_\lambda}\frac{\partial\psi^*}{\partial x_\lambda}-\sigma^2\psi\psi^*+\tfrac{1}{2}g\psi^2\psi^{*2}, \tag{1}$$

where σ^2 and g are positive constants and

$$(x_1, x_2, x_3, x_4)=(x, y, z, ict).$$

This field was recently discussed to some extent by Finkelstein, LeLevier, and Ruderman[4] incidentally to the treatment of a more complicated case.

If one uses the Lagrangian density (1) in a four-dimensional variational principle, one gets for the field equation

$$(\Box^2-\sigma^2)\psi+g\psi^2\psi^*=0. \tag{2}$$

By the usual methods one also finds that the charge-current density vector is given by

$$s_\mu=-i\epsilon(\psi\partial\psi^*/\partial x_\mu-\psi^*\partial\psi/\partial x_\mu), \tag{3}$$

where ϵ is a constant, and the energy-momentum density tensor is given by

$$T_{\mu\nu}=\frac{\partial\psi}{\partial x_\mu}\frac{\partial\psi^*}{\partial x_\nu}+\frac{\partial\psi}{\partial x_\nu}\frac{\partial\psi^*}{\partial x_\mu}+\mathfrak{L}\delta_{\mu\nu}. \tag{4}$$

2. SINGLE PARTICLE

In the stationary spherically symmetric case one sets, in polar coordinates,

$$\psi=\theta(r)e^{i\omega t}, \tag{5}$$

where θ and ω are real, and ω is constant. The field equation becomes

$$d^2\theta/dr^2+(2/r)d\theta/dr-\alpha^2\theta=-g\theta^3, \tag{6}$$

where $\alpha^2=\sigma^2-\omega^2/c^2$ will be assumed to be positive.

It follows that in this case there exist solutions that are everywhere analytic and go to zero exponentially at infinity.[5] The two simplest solutions, one without nodes, the other with one node, obtained by numerical integration, are shown in Fig. 1.[6] We shall restrict ourselves hereafter to the nodeless solution, as representing the ground state of the particle.

From the solution describing a particle, on the basis of Eqs. (3) and (4), one obtains for the charge of the particle,

$$Q\equiv-ic^{-1}\int s_4dV=2\epsilon\omega c^{-2}\int\theta^2dV, \tag{7}$$

and for the energy of the particle, after making use of the field equation and carrying out an integration by

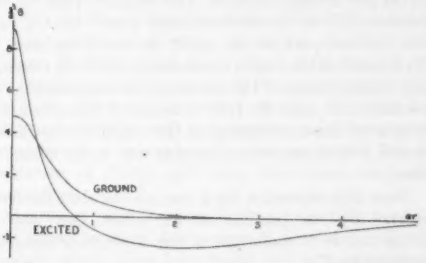


FIG. 1. The two simplest particle-like solutions of Eq. (6).

* Now at the Naval Research Laboratory, Washington, D. C.
[1] N. Rosen, Phys. Rev. **55**, 94 (1939).
[2] R. J. Finkelstein, Phys. Rev. **75**, 1079 (1949).
[3] S. D. Drell, Phys. Rev. **79**, 220 (1950).
[4] Finkelstein, LeLevier, and Ruderman, Phys. Rev. **83**, 326 (1951); henceforth referred to as FLR.
[5] This is shown by FLR by an analysis of the solutions in the phase plane. It follows from their arguments that there exists at least one particle-like solution with any given number of nodes. This result may perhaps be more directly seen by transforming (6) into $\eta''/\eta=1-(\eta/x)^2$ by $x=\alpha r$, $\eta=g^{\frac{1}{2}}r\theta$, and considering solutions with $\eta(0)=0$ as functions of increasing $\eta'(0)$.
[6] The solution exhibited in Fig. 2 of FLR is, as nearly as can be determined in the absence of a scale, identical with our nodeless solution.

parts,

$$W \equiv -\int T_{44} dV = \int (2\omega^2 c^{-2}\theta^2 + \tfrac{1}{2}g\theta^4) dV. \quad (8)$$

Thus the energy of the particle is positive definite, although the energy density $-T_{44}$ is not.

From Eq. (6) it is seen that, for large values of r for which θ is small so that the right-hand side is negligible, the particle-like solution will have a behavior given by

$$\theta = Ag^{-1}e^{-\alpha r}/r, \quad (9)$$

where A is a constant. A rough numerical integration gives $A \sim 2.5$.

It is seen that the particle-like solutions of (6) depend on the frequency ω. In the case of a neutral particle Eq. (7) shows that $\omega = 0$, so that ψ is real. For a charged particle some criterion would have to be adopted to fix the value of ω, the sign of which depends on the sign of the charge.

3. INTERACTION BETWEEN PARTICLES

We now consider the question of the interaction between two identical particles which are nearly at rest. The exact way of dealing with two particles would be to find a solution of the field equation (2) for which there are two small regions in which $|\psi|$ has an appreciable value. However, this is not feasible. Instead, we limit ourselves to the case in which the two particles are far apart and take as the approximate solution of (2),

$$\psi = e^{i\omega t}[\theta(r_A) + \theta(r_B)], \quad (10)$$

where the points A and B are the centers of the two particles, r_A and r_B are the distances from A and B to a point in the field, and $\theta(r)$ is the ground-state solution of (6) for a single particle. The assumed form of the solution (10) is justified to some extent by the fact that for two particles far apart the nonlinear term in (2) is small in the region where $\theta(r_A)$ and $\theta(r_B)$ overlap, and hence the sum of the two solutions is approximately a solution. In order for (10) to be a good approximation in spite of the nonlinearity of the equation, the points A and B must move in a suitable way in the course of time.

From this expression for ψ one can calculate the force exerted by one particle on the other. One way of doing this is to make use of the divergence equation satisfied by $T_{\mu\nu}$,

$$\partial T_{\mu\lambda}/\partial x_\lambda = 0. \quad (11)$$

In particular, if for $\mu = k = 1, 2, 3$ we integrate this equation over a three-dimensional volume V bounded by a closed surface S and make use of Gauss's theorem we get

$$\frac{d}{dx_4}\int_V T_{k4} dV = -\int_S T_{kn} dS, \quad (12)$$

where the subscript n indicates the component in the direction of the outward normal. Since the left-hand side of (12) represents the time-rate of change of momentum contained in V, the right-hand side represents the force acting on this region, which also follows from the fact that the space components of $T_{\mu\nu}$ represent stress components. Thus we can write the force components

$$F_k = -\int_S T_{kn} dS. \quad (13)$$

If we choose the surface S so that it encloses one particle (this will be practically the case if the smallest distance from the particle center to the surface is large compared to $1/\alpha$), then (13) will give the force on this particle.

If one chooses the surface S sufficiently large so that everywhere on it Eq. (9) is valid for both $\theta(r_A)$ and $\theta(r_B)$, the calculation becomes rather simple. Thus, suppose that at a certain moment A and B are located on the z-axis each a distance $R/2$ from the origin, R being the inter-particle distance, and we take for S the plane of symmetry midway between the particles (closed with an infinite hemisphere, which gives no contribution). It is enough to calculate

$$F_z = -\int_S T_{zz} dS,$$

and in the expression for T_{zz} it is enough to consider only the terms involving products of $\theta(r_A)$ and $\theta(r_B)$ or their derivatives to get the particle interaction. For a sufficiently large value of R one can neglect the fourth-degree terms in T_{zz}, since on S these will be small compared to the remaining terms. Under these conditions, making use of the symmetry of the problem, one finds that

$$T_{zz} = -4\alpha^2\theta^2(r_A), \quad (14)$$

and one gets for the force of attraction

$$F = 8\pi A^2\alpha^2 g^{-1}\int_0^\infty \exp[-2\alpha(\rho^2 + R^2/4)^{\frac{1}{2}}] \times (\rho^2 + R^2/4)^{-1}\rho d\rho \quad (15)$$

$$= -8\pi A^2\alpha^2 g^{-1}\,\mathrm{Ei}(-\alpha R),$$

or, assuming αR to be large, the force of attraction is given by

$$F(R) = 8\pi A^2\alpha g^{-1}e^{-\alpha R}/R, \quad (16)$$

and the interaction potential energy therefore by

$$U(R) = -8\pi A^2 g^{-1}e^{-\alpha R}/R. \quad (17)$$

For a neutral particle, $\omega = 0$, $\alpha = \sigma$, this gives

$$U(R) = -8\pi A^2 g^{-1}e^{-\sigma R}/R. \quad (18)$$

We see then that the interaction between the two

particles at large distances is described by the Yukawa potential.[7]

An inspection of the method used to obtain this result shows that it would also be obtained under more general conditions than those assumed here. If, in the Lagrangian density (1), the last term were replaced by some other function which also led to particle-like solutions, and which was negligible compared to the other terms at large distances from the particle center, then the same result would be obtained for the interaction, except for the numerical coefficient. For example, this would be the case if the last term were taken proportional to $(\psi\psi^*)^n$ for any $n>1$.

[7] H. Yukawa, Proc. Phys.-Math. Soc. Japan (3) **17**, 48 (1935).

PHYSICAL REVIEW VOLUME 85, NUMBER 2 JANUARY 15, 1952

Fine Structure of the Hydrogen Atom.* III

WILLIS E. LAMB, JR.†
Columbia Radiation Laboratory, Columbia University, New York, New York
(Received September 19, 1951)

The third paper of this series provides a theoretical basis for analysis of precision measurements of the fine structure of hydrogen and deuterium. It supplements the Bechert-Meixner treatment of a hydrogen atom by allowing for the presence of a magnetic field, as well as radiative corrections. The theory of hyperfine structure is somewhat extended. Stark effects due to motional and other electric fields are calculated. Possible radiative and nonradiative corrections to the shape and location of resonance peaks are discussed. Effects due to the finite size of the deuteron are also considered.

A theory of the sharp resonances $2^2S_{\frac{1}{2}}(m_s=\frac{1}{2})$ to $2^2S_{\frac{1}{2}}(m_s=-\frac{1}{2})$ is given which leads to an understanding of the peculiar shapes of resonance curves shown in Part II. In this connection, a violation of the "no-crossing" theorem of von Neumann and Wigner is exhibited for the case of decaying states.

THE earlier Parts[67] I and II of this paper have described some qualitative studies of the fine structure of hydrogen and deuterium made by a microwave method. In order to prepare the ground for analysis of much more highly precise measurements in Part IV, it is necessary to make available a more refined theory of the hydrogen atom than was used previously. The object of Part III is to supply this need, as well as to treat a number of other theoretical problems which arise in the work. Frequent references to Parts I and II are made. Chapters, sections, figures, tables, equations, and footnotes of Part III are numbered consecutively after those of Parts I and II.

J. ENERGY LEVELS OF A HYDROGEN-LIKE ATOM

48. General Program

The results of theory for the energy levels of an ideal hydrogen atom were given in Part I assuming an infinitely heavy nucleus, thereby neglecting reduced mass effects as well as magnetic and retarded interaction between electron and nucleus. In addition, a number of other approximations were made. The calculation of hyperfine structure was oversimplified by assumption of Back-Goudsmit and Russell-Saunders coupling. In the theories of Zeeman effect and doublet separation $P_{\frac{1}{2}}-P_{\frac{1}{2}}$ the anomalous magnetic moment of the electron was neglected. Shifts of levels due to Stark effect and relativistic and higher order corrections to Zeeman splitting were ignored.

There is no one place in the literature where a treatment of all these effects may be found. One may only form a patchwork Hamiltonian by collecting separate terms from papers by various authors who have been concerned with limited aspects of the problem. It would probably not be justified here to give a detailed systematic theory, but it does seem worthwhile to indicate the basis of the rather provisional treatment which is now possible. The object is to write down all terms known at present having a potential magnitude of 0.1 Mc/sec or larger in the discussion of the precision experiments of Part IV.

The electron and proton should be allowed to interact with one another through their intermediate coupling with the quantized electromagnetic field and the vacuum of occupied negative energy states for electrons and protons. By eliminating these effects from the theory, one hopes to find an equivalent two-body problem in which the two particles have a velocity and spin dependent interaction with one another, and the particles themselves have somewhat changed properties (renormalization of charge and mass, anomalous magnetic moments, etc.).

At present, this program has not been fully carried out. Those terms of low orders in the fine structure constant which have been found will be incorporated into the following discussion. It should be relatively easy to make the small corrections necessary when any missing terms have been calculated.

The starting point is here taken to be a two-body Dirac equation for electron and nucleus. Even when the nucleus is a proton and not a deuteron there might be grave doubt that it would obey an equation of the Dirac type in view of its anomalous magnetic moment.

* Work supported jointly by the Signal Corps and ONR.
† Present address: Department of Physics, Stanford University, Stanford, California.
[67] W. E. Lamb, Jr., and R. C. Retherford, Part I, Phys. Rev. **79**, 549 (1950), and Part II, **81**, 222 (1951).

If the additional magnetic moment is forced into the theory by insertion of terms[68] of the type used by Pauli, and if the calculation is not carried too far, it is possible to combine all terms referring to the magnetic moment of the proton so that only the observed moment enters the final equations. In this approximation it is reasonable to use the same result for other nuclei which do not even obey Fermi-Dirac statistics.

49. Two-Body Wave Equation

The two-body wave equation applicable to a hydrogen atom will be taken as

$$[H_1+H_2+e_1e_2/r+V+\Omega+H(\text{h.f.})]\Psi=E\Psi \quad (103)$$

where

$$H_1=\alpha_1\cdot(\mathbf{p}_1-e_1\mathbf{A}_1)+\beta_1m_1+e_1V_1, \quad \text{etc.} \quad (104)$$

where the index (1) refers to the electron and (2) to the proton. Relativistic units are used in which $\hbar$, c, and m are set equal to unity. Then $m_1=m=1$ and $m_2=M\sim1836$, and $e_1=-e$, $e_2=Ze$ while e^2 can be replaced by the fine structure constant $\alpha=e^2/hc\sim1/(137.043)$. The wave function Ψ has $4\times4=16$ components $\Psi_{n_1n_2}$ where the Dirac matrices $\boldsymbol{\alpha}_1$, β_1 act on the first index $n_1=1, 2, 3, 4$ and $\boldsymbol{\alpha}_2$, β_2 act on the second index n_2. $\mathbf{A}$ and V are the vector and scalar potentials of the external electromagnetic field.

The interaction between the particles includes the Coulomb potential energy

$$U=e_1e_2/r \quad (105)$$

as well as the Breit[69]-Darwin[70] magnetic-retarded interaction

$$V=-(e_1e_2/2r)[(\boldsymbol{\alpha}_1\cdot\boldsymbol{\alpha}_2)+(\boldsymbol{\alpha}_1\cdot\mathbf{r})(\boldsymbol{\alpha}_2\cdot\mathbf{r})/r^2]. \quad (106)$$

The finite remnants for a bound electron of the electromagnetic self-energy are represented by an operator Ω which allows for the anomalous magnetic moment of the electron and the electromagnetic level shift, etc.

The hyperfine energy term $H(\text{h.f.})$ is intended to represent only that part of the hyperfine interaction which arises from the non-Dirac part $\boldsymbol{\mu}_I'=\mu_I'\boldsymbol{\sigma}_2$ of the nuclear moment,

$$H(\text{h.f.})=-e_1\boldsymbol{\alpha}_1\cdot\mathbf{A}_I'(\mathbf{r}_1)-\boldsymbol{\mu}_I'\boldsymbol{\sigma}_2\cdot\mathbf{H}(\mathbf{r}_2) \quad (107)$$

where

$$\mathbf{A}_I'=(\boldsymbol{\mu}_I'\times\mathbf{r})/r^3. \quad (108)$$

While the first three terms in the Hamiltonian (103) should be taken literally and treated exactly, it is known that this may not be done with V which must in fact be evaluated using only first-order perturbation theory. This was established in 1929–1932 by Breit[69] in his discussion of the triplet fine structure of helium, and is connected with the omission of yet uncalculated fourth-order terms in e.

For similar reasons, or because of their smallness, the subsequent terms in the Hamiltonian will also be taken only in first order. This means that it suffices to calculate only

$$\bar{\Omega}=\int\Psi^*\Omega\Psi d\tau, \quad (109)$$

etc. where Ψ is the eigenfunction of the hydrogen atom with the Hamiltonian H_1+H_2+U without external fields. The solution of the underlying two-body Dirac equation is thereby much simplified, and subsequent calculation of the above averages is made fairly easy.

50. Reduction of Wave Equation

Two methods of approach have been used to solve the two-body Dirac equation. The first, used by Breit[69] and applied by Bechert and Meixner[71] to hydrogen fine structure, involved the reduction from a $4\times4=16$ component wave equation to a $2\times2=4$ component wave equation. Unfortunately, the treatment by Bechert and Meixner contains some errors which will be pointed out in the following discussion. More recently, Breit and Brown[72] gave a reduction to an 8 component wave equation. Both treatments led to the conclusion that the fine structure as calculated from the one-body Dirac equation was correct up to and including order α^2R except for the appearance of the reduced mass $\mu=mM/(m+M)$ in the expected manner and a common shift $\alpha^2R/(64M)$ for all $n=2$ levels.

The reduction to four components is used here because it permits a closer connection with the more elementary treatment given in Part I. Writing

$$\Psi=\begin{pmatrix}\phi\\ \omega_1\\ \omega_2\\ \chi\end{pmatrix}, \quad (110)$$

where the functions ϕ, ω_1, ω_2, and χ are four component wave functions with

$$\phi=\begin{pmatrix}\Psi_{11}\\ \Psi_{12}\\ \Psi_{21}\\ \Psi_{22}\end{pmatrix}, \quad \omega_1=\begin{pmatrix}\Psi_{31}\\ \Psi_{32}\\ \Psi_{41}\\ \Psi_{42}\end{pmatrix}, \quad \omega_2=\begin{pmatrix}\Psi_{13}\\ \Psi_{14}\\ \Psi_{23}\\ \Psi_{24}\end{pmatrix} \text{ and } \chi=\begin{pmatrix}\Psi_{33}\\ \Psi_{34}\\ \Psi_{43}\\ \Psi_{44}\end{pmatrix}. \quad (111)$$

[68] W. Pauli, Handbuch der Physik, second edition **24/1**, 233 (1933).
[69] G. Breit, Phys. Rev. **34**, 553 (1929); **39**, 616 (1932).
[70] C. G. Darwin, Phil. Mag. **39**, 537 (1920).
[71] K. Bechert and J. Meixner, Ann. Physik **22**, 525 (1935).
[72] G. Breit and G. E. Brown, Phys. Rev. **74**, 1278 (1948). Also, T. Ishidzu, Prog. Theor. Phys. **6**, 48 (1951).

we have

$$\alpha_1\Psi=\sigma_1\begin{pmatrix}\omega_1\\ \phi\\ \chi\\ \omega_2\end{pmatrix},\quad \beta_1\Psi=\begin{pmatrix}\phi\\ -\omega_1\\ \omega_2\\ -\chi\end{pmatrix},$$
$$\alpha_2\Psi=\sigma_2\begin{pmatrix}\omega_2\\ \chi\\ \phi\\ \omega_1\end{pmatrix}\quad\text{and}\quad \beta_2\Psi=\begin{pmatrix}\phi\\ \omega_1\\ -\omega_2\\ -\chi\end{pmatrix}, \tag{112}$$

where the σ's are four component spin matrices. The equation

$$[H_1+H_2+U-E]\Psi=0 \tag{113}$$

is then equivalent to the four equations

$$\begin{aligned}
&\sigma_1\cdot(\mathbf{p}_1-e_1\mathbf{A}_1)\omega_1+\sigma_2\cdot(\mathbf{p}_2-e_2\mathbf{A}_2)\omega_2\\
&\qquad+(m_1+m_2+U-E)\phi=0, \quad \text{(a)}\\
&\sigma_1\cdot(\mathbf{p}_1-e_1\mathbf{A}_1)\phi+\sigma_2\cdot(\mathbf{p}_2-e_2\mathbf{A}_2)\chi\\
&\qquad+(m_2-m_1+U-E)\omega_1=0, \quad \text{(b)}\\
&\sigma_1\cdot(\mathbf{p}_1-e_1\mathbf{A}_1)\chi+\sigma_2\cdot(\mathbf{p}_2-e_2\mathbf{A}_2)\phi\\
&\qquad+(m_1-m_2+U-E)\omega_2=0, \quad \text{(c)}\\
&\sigma_1\cdot(\mathbf{p}_1-e_1\mathbf{A}_1)\omega_2+\sigma_2\cdot(\mathbf{p}_2-e_2\mathbf{A}_2)\omega_1\\
&\qquad+(-m_1-m_2+U-E)\chi=0. \quad \text{(d)}
\end{aligned} \tag{114}$$

In reducing these equations to one for the large component ϕ, it is desirable to assign orders of magnitude in α and $1/M$ to the various terms. If the rest energy of the electron is taken as unity, that of the proton is $M\sim1836$, the Rydberg energy $hcRZ^2$ is $\frac{1}{2}\alpha^2Z^2$ and the fine structure doublet splitting for $n=2$ is $\frac{1}{16}\alpha^2Z^4R$ or $\frac{1}{32}\alpha^4Z^4\sim10{,}950$ Mc/sec for hydrogen. The radiative width of $2\dot{p}$ is $(2^9/3^8)\alpha^2Z^4hcR$ or $(2/3)^8\alpha^5Z^4\sim99.7$ Mc/sec for $Z=1$, while the electromagnetic shift[73] of the $2^2S_{\frac{1}{2}}$ level for hydrogen is

$$\delta\sim\frac{\alpha^5Z^4}{6\pi}\log(mc^2/Z^2\bar{\epsilon})\sim1040\text{ Mc/sec} \tag{115}$$

and abnormally large for its order because of the logarithm involving an atomic excitation energy $\bar{\epsilon}$. The term of order α^6 in the doublet separation, which according to Dirac's exact solution for hydrogen is $(5/256)\alpha^6\sim0.364$ Mc/sec, cannot be obtained even for the one-body problem by the method of reduction to two components, but this term may simply be borrowed from the exact treatment for $M=\infty$. Terms effectively of order α^6 due to an external magnetic field are correctly given by the subsequent treatment.

Turning now to terms in $1/M$, those of order α^2/M clearly correspond to reduced mass corrections to the Bohr energies. Any terms of orders α^3/M, α^4/M, α^5/M, and α^4/M^2 might be of importance for present microwave experiments, but calculation shows that of these only α^4/M and α^4/M^2 are actually present. Both have been examined, but the latter are numerically negligible, and for simplicity will be omitted in the subsequent discussion.

The Zeeman splitting in practice will be less than, but comparable to, the fine structure doublet splitting, so that the vector potential of the applied magnetic field will be overestimated if counted as of order given by

$$e_1\alpha_1\cdot\mathbf{A}_1\sim\alpha^4/32\quad\text{or}\quad e_1A_1\sim\alpha^3/16 \tag{116}$$

with $eH_1\sim\alpha^4/16$. The Coulomb attraction

$$U=e_1e_2/r \tag{105}$$

has an average value of order $Z^2\alpha^2/4$, but for r equal to the classical electron radius $r_0=e^2/mc^2=\alpha$, U acquires the larger order unity so that terms involving it must be treated with more care. It may be remarked in passing that hyperfine splittings are of order α^4/M.

Taking ϕ of order unity, one sees that $\omega_1\sim\alpha$, $\omega_2\sim\alpha/M$, and $\chi\sim\alpha^2/M$. If the equation for ϕ derived from (114) is to be correct to order α^5/M, it is necessary that ω_1, and ω_2 be correctly calculated to order α^4/M. According to (114b) this requires that χ be known correctly to order α^3/M. From (114d), one obtains with this accuracy

$$\chi=(2M)^{-1}(\sigma_2\cdot\mathbf{p}_2)\omega_1. \tag{117}$$

Likewise from (114c)

$$\begin{aligned}
\omega_2=(2M-U+W)^{-1}&[\sigma_1\cdot(\mathbf{p}_1-e_1\mathbf{A}_1)\chi\\
&+\sigma_2\cdot(\mathbf{p}_2-e_2\mathbf{A}_2)\phi],
\end{aligned} \tag{118}$$

where $W=E-M-1$ is the nonrelativistic energy, or to the requisite order

$$\omega_2=(2M)^{-1}\sigma_2\cdot(\mathbf{p}_2-e_2\mathbf{A}_2)\phi \tag{119}$$

neglecting for $r\sim\alpha$ terms in ω_2 of order α/M^2 which slightly exceed the stated order α^3/M. However, a more careful consideration shows this neglect to be justified because of the small volume involved. Finally

$$\omega_1=(2-U+W)^{-1}[\sigma_1\cdot(\mathbf{p}_1-e_1\mathbf{A}_1)\phi+\sigma_2\cdot\mathbf{p}_2\chi]. \tag{120}$$

Insertion of these expressions in (114a) gives

$$\begin{aligned}
&\sigma_1\cdot(\mathbf{p}_1-e_1\mathbf{A}_1)(2-U+W)^{-1}[\sigma_1\cdot(\mathbf{p}_1-e_1\mathbf{A}_1)\\
&\quad+(2M)^{-1}(\sigma_2\cdot\mathbf{p}_2)^2(2-U+W)^{-1}\sigma_1\cdot\mathbf{p}_1]\phi\\
&\quad+(2M)^{-1}\sigma_2\cdot(\mathbf{p}_2-e_2\mathbf{A}_2)\sigma_2\cdot(\mathbf{p}_2-e_2\mathbf{A}_2)\phi\\
&\qquad+(U-W)\phi=0.
\end{aligned} \tag{121}$$

After some reduction, keeping terms of order up to and including α^5/M and replacing $W-U$ by $\frac{1}{2}\mathbf{p}_1^2+\frac{1}{2}(\mathbf{p}_2^2/M)$ in terms of order α^4, one finds a wave equation for ϕ

$$(H_I+H_{II}+H_{III})\phi=W\phi \tag{122}$$

where the Hamiltonian

$$\begin{aligned}
H_I=\tfrac{1}{2}\mathbf{p}_1^2-\tfrac{1}{8}\mathbf{p}_1^4+(\nabla U\cdot\mathbf{p}_1)/4i\\
+(2-U+W)^{-2}\sigma_1\cdot\nabla_1U\times\mathbf{p}_1+U
\end{aligned} \tag{123}$$

[73] H. A. Bethe, Phys. Rev. **72**, 339 (1947).

is familiar from the reduction[74] of the one-body Dirac equation to 2-component form, and for a hydrogen-like atom with fixed point nucleus gives the fine structure correctly to order α^4. The term

$$H_{II}=\tfrac{1}{2}(\mathbf{p}_2^2/M) \tag{124}$$

gives the kinetic energy of the nucleus, while

$$H_{III}=-(2-U+W)^{-1}(2e_1\mathbf{A}_1\cdot\mathbf{p}_1+e_1\boldsymbol{\sigma}_1\cdot\mathbf{H}_1) +\tfrac{1}{2}e_1^2\mathbf{A}_1^2-(2M)^{-1}(2e_2\mathbf{A}_2\cdot\mathbf{p}_2+e_2\boldsymbol{\sigma}_2\cdot\mathbf{H}_2) +(e_1i/4)(\boldsymbol{\sigma}_1\cdot\boldsymbol{\nabla}U)(\boldsymbol{\sigma}_1\cdot\mathbf{A}_1) \tag{125}$$

gives the interaction of the atom with an external magnetic field, except for some terms from Ω and H(h.f.) which are inserted later.

The energy contributed by the Darwin-Breit term Y is given by the average

$$\bar{Y}=-\tfrac{1}{2}e_1e_2\int\begin{pmatrix}\phi\\ \omega_1\\ \omega_2\\ \chi\end{pmatrix}^*\left[\frac{\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2}{r}+\frac{\boldsymbol{\sigma}_1\cdot\mathbf{r}\boldsymbol{\sigma}_2\cdot\mathbf{r}}{r^3}\right]\begin{pmatrix}\chi\\ \omega_2\\ \omega_1\\ \phi\end{pmatrix}d\tau \tag{126}$$

which to the required order α works out to be just the average of the operator

$$H_{IV}=[e_1e_2/(4M)][(\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2)/r^3-3(\boldsymbol{\sigma}_1\cdot\mathbf{r})(\boldsymbol{\sigma}_2\cdot\mathbf{r})/r^5] -[e_1e_2/(8M)](1+\tfrac{1}{2}\alpha/r)^{-3}r^{-2}\frac{dU}{dr} \times[\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2-(\boldsymbol{\sigma}_1\cdot\mathbf{r}\boldsymbol{\sigma}_2\cdot\mathbf{r}/r^2)] -[e_1e_2/(2M)][r^{-1}\mathbf{p}_1\cdot\mathbf{p}_2+r^{-3}\mathbf{r}\cdot\mathbf{p}_1'\mathbf{r}\cdot\mathbf{p}_2'] -[e_1e_2/(2Mr^2)][\boldsymbol{\sigma}_2\cdot\mathbf{r}\times\mathbf{p}_1-\boldsymbol{\sigma}_1\cdot\mathbf{r}\times\mathbf{p}_2] \tag{127}$$

for the state whose wave function is ϕ. The prime on $\mathbf{p}'$ indicates that the operator does not act on $\mathbf{r}$.

51. Exhibition of Reduced Mass

Bechert and Meixner[71] showed that when $\mathbf{p}_1=-\mathbf{p}_2=\mathbf{p}$ the Hamiltonian

$$H_I+H_{II}+H_{IV}$$

could be rearranged so as to exhibit explicitly the reduced mass μ by writing

$$H_I+H_{II}+H_{IV}=H_a+H_b+H_c, \tag{128}$$

where

$$H_a=\mathbf{p}^2/(2\mu)-\mathbf{p}^4/(8\mu^3)+(\boldsymbol{\nabla}U\cdot\mathbf{p})/(4i\mu^2) +\mu^{-2}(2-U+W)^{-3}\boldsymbol{\sigma}_1\cdot\boldsymbol{\nabla}U\times\mathbf{p}+U \tag{129}$$

is the Hamiltonian of type (123) for a particle of mass μ.

[74] See, for example, L. I. Schiff, *Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1949), p. 320.

52. Fine and Hyperfine Energies

The remaining terms from

$$H_I+H_{II}+H_{IV}$$

may be written as

$$H_b=(3/8M)\mathbf{p}^4-(\boldsymbol{\nabla}U\cdot\mathbf{p})/(2iM) +(e_1e_2/2M)(r^{-1}\mathbf{p}^2+r^{-3}\mathbf{r}\cdot\mathbf{p}'\mathbf{r}\cdot\mathbf{p}') \tag{130}$$

and

$$H_c=(e_1e_2/4M)[r^{-3}\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2-3r^{-5}\boldsymbol{\sigma}_1\cdot\mathbf{r}\boldsymbol{\sigma}_2\cdot\mathbf{r}] -[e_1e_2/(2Mr)](\boldsymbol{\sigma}_2\cdot\mathbf{r}\times\mathbf{p}_1)-[e_1e_2/(8M)] \times(r+\tfrac{1}{2}\alpha)^{-2}\frac{dU}{dr}(\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2-r^{-2}\boldsymbol{\sigma}_1\cdot\mathbf{r}\boldsymbol{\sigma}_2\cdot\mathbf{r}). \tag{131}$$

H_c gives contributions to the hyperfine energy and will be considered further in Sec. 56.

The above equations are equivalent to those used by Bechert and Meixner, except that their terms containing

$$\boldsymbol{\nabla}U\cdot\mathbf{p}$$

in H_a and H_b have the wrong sign. These errors are compensated by their use of incorrect expressions for the averages of $\mathbf{p}^4$

$$(r^{-1}p^2+r^{-3}\mathbf{r}\cdot\mathbf{p}'\mathbf{r}\cdot\mathbf{p}')$$

which in fact have the values for the nl state of hydrogen of

$$(\mathbf{p}^4)_{kl}=(-3+4n/(l+\tfrac{1}{2}))(\alpha^4/n^4) \tag{132a}$$

$$(r^{-1}\mathbf{p}^2+r^{-3}\mathbf{r}\cdot\mathbf{p}'\mathbf{r}\cdot\mathbf{p}')_{kl} =(-2+3n/(l+\tfrac{1}{2})-2n\delta_{l0})(\alpha^3/n^4). \tag{132b}$$

Also

$$(\boldsymbol{\nabla}U\cdot\mathbf{p})/i=e_1e_2(r^{-2}\partial/\partial r)_{kl}=(2\alpha^4/n^3)\delta_{l0}. \tag{132c}$$

The discrepancies between these values and those of Bechert and Meixner arise because the results depend on whether or not a small sphere about $r=0$ is excluded. In the above equations, care has been taken to find the correct interpretation for the singularities. No sphere is to be excluded in (132a), but must be for (132b). The result is then that H_b has the same average value

$$-\alpha^4/(8Mn^4)$$

for all the fine structure levels for a given n, and therefore the separations predicted by H_a are not disturbed. Although this result was derived by treating the nucleus as a particle obeying Dirac's equation, it is clearly more general, and should follow to order α^4/M from any relativistic treatment of the two-body problem. The first two terms of H_b enter to compensate for the forcing of the reduced mass μ into H_a in the manner in which it would appear if the one-body Dirac equation for a particle of mass μ were reduced to two-component form. This is done merely for convenience in calculation. The last term in H_b is just the Darwin retarded-magnetic

interaction energy which has a classical analog for any distribution of currents.

In the following sections are given the detailed expressions which make up the working Hamiltonian for analysis of the experiments. These include: unperturbed energies, Zeeman, hyperfine, and Stark energies.

K. WORKING HAMILTONIAN

53. Unperturbed Levels

In the absence of an external electric field (motional or otherwise), there are no terms in the Hamiltonian which mix S and P states. This is true both for hyperfine interaction and electromagnetic term shift Ω. The unperturbed states in the absence of hyperfine and Zeeman splitting may be specified by quantum numbers n, j, m_j for the electron which are indicated in Table II. with energies 0, δ, and ΔE shown in Fig. 47. Quantum numbers I and m_I for the nuclear spin will also be used.

The position of $2^2P_{\frac{1}{2}}$ is taken as the zero of energy for convenience despite the fact that the absolute position of $2^2P_{\frac{1}{2}}$ is changed by the quantum electrodynamic effects. According to Dirac's treatment, $\delta=0$ and with inclusion of the reduced mass where it is significant

$$\Delta E=(1/32)\mu\alpha^4+(5/256)\alpha^6. \tag{133}$$

The electrodynamic shifts of levels δE_{nlj} which arise from Ω may be written as[75,76]

$$\begin{aligned}\delta E_{20\frac{1}{2}}&=(\alpha^5/6\pi)[\log(1/2k_0)+19/30],\\ \delta E_{21\frac{1}{2}}&=(\alpha^5/6\pi)[\log(\alpha^2/2k_1)-1/8],\\ \delta E_{21\frac{3}{2}}&=(\alpha^5/6\pi)[\log(\alpha^2/2k_1)+1/16],\end{aligned} \tag{134}$$

where k_0 and k_1 are average atomic excitation energies. One gets thereby an additional contribution to ΔE of $\alpha^5/(32\pi)$ or 25.40 Mc/sec due to the spin-orbit interaction association with the second-order anomalous magnetic moment[77] $(\alpha/2\pi)\mu_0$ of the electron. When the 1.38 percent fourth-order[78] reduction in this moment is included, ΔE becomes

$$\begin{aligned}\Delta E=(1/32)\mu\alpha^4+(5/256)\alpha^6\\ +(\alpha^5/32\pi)(1-5.946(\alpha/\pi)).\end{aligned} \tag{135}$$

Using the 1951 constants of Bearden and Watts[79] $\Delta E/h$ is

$$\Delta E/h=\begin{cases}10{,}967.463\text{ Mc/sec for }H\\ 10{,}970.447\text{ Mc/sec for }D.\end{cases} \tag{136}$$

The relative level shift δ according to Eq. (134) is

$$\delta=(\alpha^5/6\pi)\left[\log\left(\frac{k_1}{\alpha^2k_0}\right)+\frac{19}{30}+\frac{1}{8}\right] \tag{137}$$

[75] Bethe, Brown, and Stehn, Phys. Rev. **77**, 370 (1950).
[76] E. R. Cohen, report to be published.
[77] J. Schwinger, Phys. Rev. **73**, 416 (1948).
[78] R. Karplus and N. M. Kroll, Phys. Rev. **77**, 536 (1950).
[79] J. A. Bearden and H. M. Watts, Phys. Rev. **81**, 73 (1951).

TABLE II. Unperturbed energy levels.

	n	l	j	m_j	Energy
$2^2S_{\frac{1}{2}}$	2	0	$\frac{1}{2}$	$\pm\frac{1}{2}$	$E^0(2^2S_{\frac{1}{2}})=\delta$
$2^2P_{\frac{1}{2}}$	2	1	$\frac{1}{2}$	$\pm\frac{1}{2}$	$E^0(2^2P_{\frac{1}{2}})=0$
$2^2P_{\frac{3}{2}}$	2	1	$\frac{3}{2}$	$\pm\frac{3}{2}$, $\pm\frac{1}{2}$	$E^0(2^2P_{\frac{3}{2}})=\Delta E$

which should be 1051.41 Mc/sec for hydrogen according to Bethe, Brown and Stehn.[75] The second-order anomalous electron magnetic moment contributes 67.77 Mc/sec of this. When the effect of the subsequently calculated fourth-order magnetic moment is included, the shift becomes 1050.47 Mc/sec. According to the above authors,[75] the shift ought to be proportional[80] to the electronic reduced mass μ, which would make it 0.29 Mc/sec higher for deuterium.

There appears to be an unexplained discrepancy between this value and the previously reported[67] value for δ of 1062±5 Mc/sec. In any case, the object of the present research is to determine an experimental value for δ, and the reduction of data is entirely independent of its theoretical value.

54. Magnetic Energy

The contribution of H_{III} is given by

$$(H_{III})_{\mathrm{av}}=\int\phi^*H_{III}\phi d\tau \tag{138}$$

in which terms smaller than the usual Zeeman energy by factors $1/M$ and α^2 are to be kept. The latter requires that departures of ϕ of order α^2 from the usual nonrelativistic Schrödinger wave function be considered. These might be of two kinds: (1) a mixing of other states and (2) a renormalization. Fortunately only the latter gives a contribution. One has

$$\int(|\phi|^2+|\omega_1|^2)d\tau\sim 1$$

or

$$\int|\phi|^2d\tau\sim 1-\tfrac{1}{4}(\rho^2)_{\mathrm{av}}=1+\tfrac{1}{2}W \tag{139}$$

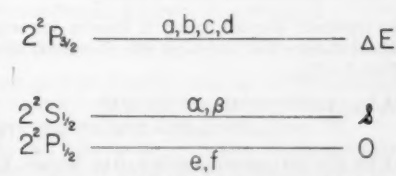


FIG. 47. Energy levels for $2^2S_{\frac{1}{2}}$ in zero magnetic field without hyperfine structure. The magnetic sub-levels are indicated by letters α, β, a, b, c, d, e, and f as in Fig. 14.

[80] Bethe, Brown, and Stehn did not consider radiative processes involving the nucleus. Unpublished calculations by B. S. Gourary which take these into account indicate a somewhat larger dependence on reduced mass.

TABLE III. Coefficients of Eq. (143) for various states.

State	α	β	a	b	c	d	e	f
Coefficient	7/36	7/36	6/36	4/36	4/36	6/36	5/36	5/36

where the nonrelativistic energy is

$$W=-\alpha^2/8 \tag{140}$$

for $n=2$. Then H_{III} is equivalent to the average of the operator

$$\begin{aligned}H_{III}\rightarrow&-\tfrac{1}{2}(2e_1\mathbf{A}_1\cdot\mathbf{p}_1+e_1\boldsymbol{\sigma}_1\cdot\mathbf{H}_1)\\&-\tfrac{1}{2}(2e_2\mathbf{A}_2\cdot\mathbf{p}_2+e_2\boldsymbol{\sigma}_2\cdot\mathbf{H}_2)/M\\&+\tfrac{1}{2}e_1^2\mathbf{A}_1^2+W\mu_0(\mathbf{L}+2\mathbf{S})\cdot\mathbf{H}\\&+\tfrac{1}{8}(\boldsymbol{\sigma}_1\cdot\nabla U)\cdot(\mathbf{H}\times\mathbf{r})\end{aligned} \tag{141}$$

calculated with the usual Schrödinger wave function, where $\mu_0=\frac{1}{2}(\alpha)^1$ is the Bohr magneton. The term $e_2\boldsymbol{\sigma}_2\cdot\mathbf{H}/8M$ will be treated with other contributions to hyperfine energy in Sec. 56.

(a) *Relativistic Corrections to Magnetic Moment*

The last two terms of (141), in case of Russell-Saunders coupling, are equivalent to the relativistic corrections of order α^2 to the magnetic moment of the atom as first calculated by Breit.[81] According to these, the effective Landé g-values for the various states are given by

$$g_j=g_j^{(0)}\begin{cases}1+\frac{2}{3}W & 2^2S_{\frac{1}{2}}\\ 1+2W & 2^2P_{\frac{1}{2}},\\ 1+\frac{4}{5}W & 2^2P_{\frac{3}{2}}\end{cases} \tag{142}$$

where

$$g_j^{(0)}=2,\ \tfrac{2}{3},\text{ and }4/3 \tag{142a}$$

for these states, respectively. For a magnetic field of 1500 gauss, larger than those used in the precision determinations of Part IV, the correction to the frequency of transition $\alpha\beta$ amounts to only 0.02 Mc/sec, and is no larger for any other transition. The relativistic moment corrections may therefore be neglected, although they are of the retained order α^6.

(b) *Quadratic Zeeman Energy*

The quadratic Zeeman energy is likewise too small to require a correction, although also of nominal order α^6. Thus for $2^2S_{\frac{1}{2}}$

$$\begin{aligned}(\tfrac{1}{2}e_1^2\mathbf{A}_1^2)_{\text{Av}}&=\tfrac{1}{8}e^2H^2(r^2\sin^2\theta)_{\text{Av}}=(7/2)e^2a_0^2H^2\\&=7(\mu_0H)^2/(hcR)=(7/36)\alpha^2x^2\Delta E\end{aligned} \tag{143}$$

which for $H=1500$ gauss amounts to 0.01 Mc/sec. The values of the coefficient of $\alpha^2x^2\Delta E$ in Eq. (143) for the other Zeeman components, assuming Russell-Saunders coupling, are given in Table III.

[81] G. Breit, Nature **122**, 649 (1928), also Mott and Massey, *Theory of Atomic Collisions* (Oxford University Press, London, 1933), first edition, pp. 47–57; H. Margenau, Phys. Rev. **57**, 383 (1940).

(c) *Energy of Orientation of Electron Spin*

To the term $-\frac{1}{2}e_1\boldsymbol{\sigma}_1\cdot\mathbf{H}_1$ in (141) must be added the contribution from Ω of the anomalous magnetic moment of the electron, giving in usual units an orientation energy

$$g_S\mu_0\mathbf{S}\cdot\mathbf{H} \tag{144}$$

where the Landé g-factor is[77,78]

$$g_S=2(1+(\alpha/2\pi)-2.973\alpha^2/\pi^2). \tag{145}$$

(d) *Reduced Mass Effects and Motional Electric Field*

The remaining terms of H_{III} are

$$H_{III}'=-(e_1/m_1)\mathbf{A}_1\cdot\mathbf{p}_1-(e_2/m_2)\mathbf{A}_2\cdot\mathbf{p}_2. \tag{146}$$

It will be shown how these terms imply the introduction of an effective g value for orbital motion

$$g_L=1-(1/M) \tag{146a}$$

and also contribute an electric field due to motion of the atom through a uniform magnetic field. The vector potentials are then

$$\mathbf{A}_1=\tfrac{1}{2}\mathbf{H}\times\mathbf{r}_1,\quad \mathbf{A}_2=\tfrac{1}{2}\mathbf{H}\times\mathbf{r}_2 \tag{147}$$

so that (146) becomes

$$\begin{aligned}H_{III}'=-\tfrac{1}{2}[&(e_1/m_1)(\mathbf{H}\cdot\mathbf{r}_1\times\mathbf{p}_1)\\&+(e_2/m_2)(\mathbf{H}\cdot\mathbf{r}_2\times\mathbf{p}_2)]\end{aligned} \tag{148}$$

and upon transformation to relative and center-of-mass coordinates by the equations

$$\begin{aligned}&\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2,\quad \mathfrak{M}\mathbf{R}=m_1\mathbf{r}_1+m_2\mathbf{r}_2,\quad \mathfrak{M}=m_1+m_2\\&\mathbf{p}_1=(m_1/\mathfrak{M})\mathbf{P}+\mathbf{p}\quad \mathbf{p}_2=(m_2/\mathfrak{M})\mathbf{P}-\mathbf{p}\end{aligned} \tag{149}$$

becomes

$$\begin{aligned}H_{III}'=&-(2\mathfrak{M})^{-1}(e_1+e_2)[\mathbf{H}\cdot\mathbf{R}\times\mathbf{P}]\\&-\tfrac{1}{2}(e_1m_1^{-1}-e_2m_2^{-1})(\mathbf{H}\cdot\mathbf{R}\times\mathbf{p})\\&-\tfrac{1}{2}(e_1m_2-e_2m_1)\mathfrak{M}^{-2}(\mathbf{H}\cdot\mathbf{r}\times\mathbf{P})\\&-\tfrac{1}{2}(e_1m_2m_1^{-1}+e_2m_1m_2^{-1})\mathfrak{M}^{-1}(\mathbf{H}\cdot\mathbf{r}\times\mathbf{p}).\end{aligned} \tag{150}$$

The first term is zero for a neutral atom with $e_1=-e_2$, and the last term in usual units is

$$g_L\mu_0\mathbf{H}\cdot\mathbf{L}, \tag{151}$$

where

$$\mathbf{L}=\mathbf{r}\times\mathbf{p} \tag{152}$$

is the orbital angular momentum operator, and

$$g_L=m_1^{-1}-m_2^{-1}=1-(1/M) \tag{153}$$

as stated above. Needless to say, this result can also be obtained by more elementary methods.

The third term of (150) may be written as

$$-\tfrac{1}{2}e_1(\mathbf{V}\times\mathbf{H}\cdot\mathbf{r}), \tag{154}$$

where

$$\mathbf{V}=\mathbf{P}/\mathfrak{M} \tag{155}$$

is the velocity of the atom, and represents just half of the expected Stark perturbation due to a motional electric field

$$\mathbf{E}=(\mathbf{V}\times\mathbf{H})/c. \tag{156}$$

This defect is made up by the second term of (150) which can be written as

$$-(e_1/\mu)\mathfrak{A}\cdot\mathbf{p} \tag{157}$$

with

$$\mathfrak{A}=\tfrac{1}{2}\mathbf{H}\times\mathbf{R} \tag{158}$$

the vector potential at the atomic center of mass $\mathbf{R}$. As far as the internal motion alone is concerned, this term is equivalent to a gauge transformation. Consider the wave equation for the atom

$$[\mathbf{P}^2/(2\mathfrak{M})+\mathbf{p}^2/(2\mu)-(e_1/\mu)\mathfrak{A}\cdot\mathbf{p}+g_L\mu_0\mathbf{L}\cdot\mathbf{H} -\tfrac{1}{2}e_1\mathbf{H}\cdot\mathbf{r}\times(\mathbf{P}/\mathfrak{M})+\cdots-W]\psi_1=0. \tag{159}$$

The gauge term containing

$$\mathfrak{A}=\mathrm{grad}_r(\mathfrak{A}\cdot\mathbf{r}) \tag{160}$$

may be removed by the transformation

$$\psi_1=\psi\exp(e_1 i/2\hbar)[\mathbf{H}\times\mathbf{R}\cdot\mathbf{r}]. \tag{161}$$

The operator $\mathbf{p}^2/2\mu$ acting upon ψ_1 gives an extra term which just cancels out the gauge term, while $\mathbf{P}^2/2\mathfrak{M}$ gives an additional contribution which doubles the motional Stark term. To first order in H, the result is

$$[\mathbf{P}^2/(2\mathfrak{M})+\mathbf{p}^2/(2\mu)-(e_1/\mathfrak{M})(\mathbf{P}\times\mathbf{H}\cdot\mathbf{r}) +g_L\mu_0\mathbf{H}\cdot\mathbf{L}+\cdots-W]\psi=0 \tag{162}$$

leading to the correct motional Stark effect.

55. Zeeman Energy

To summarize the results thus far, the Zeeman splitting is determined by the Hamiltonian

$$\mathcal{S}\delta_{l0}+\tfrac{2}{3}\Delta E(\mathbf{L}\cdot\mathbf{S}+\delta_{l1})+g_S\mu_0\mathbf{S}\cdot\mathbf{H}+g_L\mu_0\mathbf{L}\cdot\mathbf{H} \tag{163}$$

and the energy levels for the states $nljm_j$ may be written in a form similar to that given in Sec. 15

$$y_{a,\beta}=y_0\pm x_a, \tag{164}$$

$$y_{a,d}=\tfrac{3}{2}\pm 2x_a, \tag{165}$$

$$y_{b,c}=\tfrac{3}{4}\pm\tfrac{1}{2}x+\tfrac{1}{2}(x^2\pm x+(9/4))^{\frac{1}{2}}\pm\tfrac{1}{2}Tx, \tag{166}$$

$$y_{e,f}=\tfrac{3}{4}\pm\tfrac{1}{2}x-\tfrac{1}{2}(x^2\pm x+(9/4))^{\frac{1}{2}}\pm\tfrac{1}{2}Tx \tag{167}$$

with

$$T=(2g_L-g_S)/(g_S-g_L), \tag{168}$$

where the unit of y for all states is

$$f_1=\tfrac{2}{3}(\Delta E/h) \tag{169}$$

and

$$\mathcal{S}=f_1 y_0$$

while the units for the x's are given by

$$\text{Unit field for } x_a\text{: } H_{1a}=(2/g_S)(\tfrac{2}{3}\Delta E/\mu_0), \tag{170}$$

$$\text{Unit field for } x_\beta\text{: } H_{1\beta}=[2/(g_L+\tfrac{1}{2}g_S)](\tfrac{2}{3}\Delta E/\mu_0), \tag{171}$$

$$\text{Unit field for } x\text{: } H_1=[1/(g_S-g_L)](\tfrac{2}{3}\Delta E/\mu_0). \tag{172}$$

The values of these and other needed units using 1951 constants of Bearden and Watts[79] are given in Table IV.

In Eqs. (164)–(172) both $\mathcal{S}$ and ΔE are to be determined from observations of resonance peaks. This is inconvenient since the units of magnetic field depend on $\Delta E/\mu_0$. Fortunately the experimental value determined for $\mathcal{S}$ is highly independent of the values used for ΔE, μ_0, and h (strictly so except for nonlinearity of the Zeeman energy curves), so the procedure will be to regard $\Delta E/h$ as given by Eq. (136), and to determine $\mathcal{S}$ from the low frequency transitions ae and af. Then the high frequency transitions aa, ab, ac may be used to determine a value for the fine structure constant, assuming validity of Eq. (136). If necessary the process can be iterated should the originally assumed value of ΔE prove to be in error.

56. Hyperfine Energy

The operator for the hyperfine energy is

$$w=g_I\mu_0^2(1+\tfrac{1}{2}(\alpha/r))r^{-3}[2\mathbf{I}\cdot\mathbf{L}-g_S(\mathbf{I}\cdot\mathbf{S}-3\mathbf{I}\cdot\mathbf{r}\mathbf{S}\cdot\mathbf{r}/r^2)] +2g_S g_I\mu_0^2(r+\tfrac{1}{2}\alpha)^{-3}(dU/dr) \times[\mathbf{I}\cdot\mathbf{S}-\mathbf{I}\cdot\mathbf{r}\mathbf{S}\cdot\mathbf{r}/r^2] \tag{173}$$

which agrees with (131) except for insertion of the observed g-values for nuclear and electronic spin. Such corrections arise from terms in Ω giving the interaction of anomalous part of the magnetic moment of the electron with the nuclear magnetic moment. In evaluation of this, as usual, S states require a special consideration. There are no matrix elements of w connecting S and P states, and it therefore suffices to consider only diagonal elements for a given l value w_{ll} which have been calculated by Bethe.[82]

(a) *S States*

For $l=0$, w_{00} reduces to Eq. (48) with an additional factor $g_S/2$, if some terms of relative order α^2 are neglected. In the presence of a magnetic field, the effective Zeeman-hyperfine Hamiltonian is

$$\mathcal{H}=\mathcal{S}+g_S\mu_0\mathbf{S}\cdot\mathbf{H}+\Delta w(I+\tfrac{1}{2})^{-1}\mathbf{I}\cdot\mathbf{S}-g_S\mu_0\mathbf{I}\cdot\mathbf{H} \tag{174}$$

where Δw is the hyperfine splitting for $2^2S_{\frac{1}{2}}$ produced by the nucleus in question (taken to be $\frac{1}{2}$ the corresponding value measured[83] for $1^2S_{\frac{1}{2}}$) given in Table IV.

In Part I the energy levels were calculated for strong fields. While the exact energy levels are given by the Breit-Rabi[51] formula, it suffices for all applications here to use the next term in its high field expansion. Accordingly, to the energies given by (164)–(167) must

[82] H. A. Bethe, Handbuch der Physik, second edition **24/1**, 386 (1933).
[83] A. G. Prodell and P. Kusch, Phys. Rev. **79**, 1009 (1950).

TABLE IV. Constants used for calculations of Zeeman and hyperfine splittings.

Quantity	Hydrogen	Deuterium
f_1 (Mc/sec)	7311.642	7313.631
H_1 (gauss)	5217.801	5219.220
H_{1a} (gauss)	5222.208	5222.919
H_1 (gauss)	5209.008	5211.840
T	−0.0033703	−0.0028284
Δw (Mc/sec)	177.551	40.923

be added

$$w=\Delta w(I+\tfrac{1}{2})^{-1}m_I m_S-g_I\mu_0 H m_I +(\Delta w)^2(I+\tfrac{1}{2})^{-2}(2g_S\mu_0 H)^{-1} \times[\{I(I+1)-m_I^2\}m_S-\tfrac{1}{2}m_I]. \quad (175)$$

The last term may be written as

$$C(m_S, I, m_I)(\Delta w/g_S\mu_0 H)\Delta w \quad (176)$$

where the coefficients $C(m_S, I, m_S)$ are given in Table V.

(b) P States

For p states, the interaction (173) may be simplified to

$$w_{11}=g_I\mu_0^2\langle r^{-3}\rangle_{\text{Av}}[2\mathbf{I}\cdot\mathbf{L}+\tfrac{1}{3}g_S\{4\mathbf{I}\cdot\mathbf{S}-3\mathbf{I}\cdot\mathbf{L}\mathbf{L}\cdot\mathbf{S} -3\mathbf{L}\cdot\mathbf{S}\mathbf{I}\cdot\mathbf{L}\}] \quad (177)$$

with neglect of some terms of relative order α^2. This operator has matrix elements diagonal in $\mathbf{J}$ which can be written in the form (53) if the small difference between g_S and 2 is neglected (see Appendix VI) or as

$$(J|w_{11}|J)=\Delta w(\mathbf{I}\cdot\mathbf{J})/[2(2I+1)J(J+1)] \quad (178)$$

in terms of the hyperfine separation Δw for $2^2S_{\frac{1}{2}}$. In case of good Russell-Saunders coupling, the hyperfine splitting would be just as given by (53) or in the Back-Goudsmit limit

$$w=\tfrac{1}{4}[\Delta w/((I+\tfrac{1}{2})J(J+1))]m_I m_J. \quad (178a)$$

Since the experiments are conducted in an appreciable magnetic field, the vectors $\mathbf{L}$ and $\mathbf{S}$ are somewhat decoupled, and a correction must be applied in Eq. (178a). The elements of w_{11} off-diagonal in $\mathbf{J}$ may be written in the form

$$(J'|w_{11}|J)=\tfrac{1}{16}\Delta w(I+\tfrac{1}{2})^{-1}(J'|\mathbf{I}\cdot\mathbf{L}|J) \quad (179)$$

TABLE V. Values of coefficients $C(m_S, I, m_I)$ of Eq. (176).

		α	β
$I=\frac{1}{2}$	$m_I=\frac{1}{2}$	0	−1/4
	$m_I=-\frac{1}{2}$	1/4	0
$I=1$	$m_I=1$	0	−2/9
	$m_I=0$	2/9	−2/9
	$m_I=-1$	2/9	0

and the necessary matrix elements of $\mathbf{I}\cdot\mathbf{L}$ calculated from equations given by Condon and Shortley.[84]

The complete effective Hamiltonian for Zeeman and hyperfine energy of the $2p$ states is then

$$\mathcal{H}=\Delta E\delta_{J\frac{1}{2}}+\mu_0\mathbf{H}\cdot(g_S\mathbf{S}+g_L\mathbf{L})+w_{11}-g_I\mu_0\mathbf{H}\cdot\mathbf{I}. \quad (180)$$

In the n, J, m_J, m_I representation only the first and last terms in $\mathcal{H}$ are diagonal. If the last two terms were neglected and perturbation theory correct to all orders in $\mu_0 H/\Delta E$ were applied to the second term in (180), the result would be equivalent to a power series expansion in x of Eqs. (164–(167)). It is convenient to take the diagonal elements in J of the second term, namely

$$g_J\mu_0\mathbf{H}\cdot\mathbf{J} \quad (181)$$

where g_J is the Landé factor for state J, as part of the unperturbed Hamiltonian. The application of perturbation theory to the remainder of $\mathcal{H}$ then gives the above power series expansion as well as some new contributions arising from cross products of the second and third terms as well as powers of the third term. These contributions are of order

$$(\Delta w)^2/\mu_0 H, \quad \Delta w\mu_0 H/\Delta E, \quad (\Delta w)^2/\Delta E$$

etc., and from a rough estimate of their magnitude it is clear that only the first two need be retained for the analysis of the present experiments.

The cross-product term from second-order perturbation theory for $2^2P_{\frac{1}{2}}$ and $2^2P_{\frac{3}{2}}$ respectively may be written as

$$\pm(2/9)[g_J g_I\mu_0^2\langle r^{-3}\rangle_{\text{Av}}/\Delta E]\mu_0 H m_I \quad (182)$$

or in terms of the hyperfine splitting Δw for $2^2S_{\frac{1}{2}}$ as

$$\pm(1/36)(\Delta w/\Delta E)(I+\tfrac{1}{2})^{-1}\mu_0 H m_I. \quad (183)$$

This can be regarded as equivalent to a change in the nuclear g-value from g_I to

$$g_I'=g_I\mp(1/36)(\Delta w/\Delta E)(I+\tfrac{1}{2})^{-1}$$

as far as the energy of orientation in a magnetic field is concerned.

If it were not for this correction to g_I for p-states, the nuclear magnetic orientation energy

$$-g_I\mu_0 H m_I$$

could be ignored completely, because m_I does not change in the allowed transitions. Because of the change in effective g_I for p-states, however, there is a slight change in the separation of the resonance peaks amounting to an increase of

$$(1/18)[I/(I+\tfrac{1}{2})](\Delta w/\Delta E)\mu_0 H$$

or 0.63 and 0.19 Mc/sec/kilogauss for hydrogen and deuterium, respectively, in the case of transitions αa, αb, αc, and a decrease by the same amount for αe and αf.

[84] Condon and Shortley, *Theory of Atomic Spectra* (Cambridge University Press, London, 1935), p. 64.

This will make itself felt primarily in the observed width of a composite resonance curve and the degree of resolution of the constituent peaks for the various m_I values, but not in the apparent center of the composite curve.

The term of order

$$(\Delta w)^2/\mu_0 H$$

corresponds to the expansion for high magnetic field of the hyperfine energy given by the Breit-Rabi formula, and in fact is more general, since it applies for p states when neither I nor J is $\frac{1}{2}$. It is

$$(1/64)[\Delta w/(I+\tfrac{1}{2})]^2[J(J+1)]^{-2}(g_J\mu_0 H)^{-1} \cdot[m_J\{I(I+1)-m_I^2\}-m_I\{J(J+1)-m_J^2\}]. \quad (184)$$

57. Stark Effect

The presence of an electric field in the rf interaction space gives rise to displacements of the fine structure energy levels. Except for cases of near degeneracy such as occur in study of the $\alpha\beta$ transitions at 575 gauss for magnetic field calibration and treated in Chapter N, it suffices to use second-order perturbation theory to calculate the Stark shift of a level i

$$\Delta W_i=-\sum_n|(n|e\mathbf{E}\cdot\mathbf{r}|i)|^2/(E_n-E_i). \quad (185)$$

Assuming Russell-Saunders coupling, the matrix elements of x, y, and z in units of the Bohr radius a_0 are given in Table VI for the various transitions Although at the actually used magnetic fields, there is some departure from Russell-Saunders coupling, the Stark perturbation is small and the above matrix elements are sufficiently accurate. In addition, the denominator in Eq. (185) may be evaluated assuming a nominal level shift of 1060 Mc/sec, and a magnetic splitting given in weak field approximation by

$$g_J^{(0)}\mu_0 H m_J \quad (186)$$

using the usual Landé g-values (142a), and neglecting hyperfine splitting.

Taking only the motional electric field

$$E_y=VH/c \quad (187)$$

into account, one finds

$$|(n|e\mathbf{E}\cdot\mathbf{r}|i)|^2=e^2H^2(V^2/c^2)_{\text{Av}}|(n|y|i)|^2. \quad (188)$$

The average shift depends on $(V^2)_{\text{Av}}$ for the beam, and assuming the distribution (88)

$$V^2=(V^2)_{\text{Av}}=2U^2=2(2kT/M) \quad (189)$$

where M is the atomic mass. For an oven temperature of 2500°K, and a magnetic field $H=1159$ gauss (transition $\alpha\varepsilon$ at 2195 Mc/sec), one finds

$$(E_y)_{\text{rms}}=VH/c=10.5 \text{ volt/cm}, \quad (190)$$

TABLE VI. Values of matrix elements of coordinates for various transitions in weak magnetic field.

Transitions		x	y	z
αa	βd	$3/\sqrt{2}$	$\mp 3i/\sqrt{2}$	0
αb	βc	0	0	$\sqrt{6}$
αc	βb	$\sqrt{3}/\sqrt{2}$	$\mp i\sqrt{3}/\sqrt{2}$	0
αd	βa	0	0	0
αe	βf	0	0	$\pm\sqrt{3}$
αf	βe	$\sqrt{3}$	$\pm i\sqrt{3}$	0

and

$$eVHa_0/hc=13.47 \text{ Mc/sec} \quad (191)$$

for hydrogen, and $1/2^{\frac{1}{3}}$ as much for deuterium.

When the Stark effect shifts are small, one may neglect any asymmetries produced by them, and merely calculate their magnitude at the center of the observed resonance curve. Thus for state α of hydrogen at 1159 gauss the contributions from states a, c, and f to (185) are -0.07, -0.04, and 0.17 Mc/sec, respectively, so that state is raised by 0.06 Mc/sec. On the other hand, state e is more strongly repelled by the nearby state β and raised by an amount 0.52 Mc/sec. Accordingly the frequency for αe is lowered by 0.46 Mc/sec, which implies an increase by 0.46 Mc/sec in the level shift as calculated from the data for hydrogen without Stark effect, and by half that amount for deuterium. Similarly, there are corrections of -0.13 and -0.06 Mc/sec to be applied to results obtained from transition αf at 2395 Mc/sec. Such corrections will be applied in Part IV in the analysis of the data.

There is also an error in magnetic field calibration using transition $\alpha\beta$ at 575 gauss. Presumably, the Stark shift of state β by the degenerate level e is very small as indicated in Sec. (72), but there are shifts due to distant levels. For deuterium, these increase the frequency of $\alpha\beta$ by 0.032 Mc/sec, and therefore all magnetic fields should be lowered by a negligible fractional amount 0.032/((2.803)(5751)) or 0.02 percent.

There is evidence, discussed in Part IV, that electric fields amounting to as much as one volt per centimeter due to contact potentials or charged insulating films may be present in the interaction space. It can be seen from the above estimates that such fields would produce a wholly negligible Stark effect. The possibility of shifts produced by the rf fields used for the measurements will be discussed in Sec. 66.

58. Summary

The formulas and constants necessary for energy level calculations to the required accuracy have been given in Secs. 53–57, especially in Tables II, IV, and V and Eqs. (135), (136), (164)–(172), (175), (176), (178a), (182)–(185). A few additional effects are studied in Appendix VI, but, except for an additive contribution to δ from the finite size of the deuteron, they are numerically negligible.

L. RADIATIVE CORRECTIONS TO RESONANCE LINE FORM

59. Ideal Weisskopf-Wigner Line Form

The goal of the Weisskopf-Wigner[85] theory of radiation broadening was to obtain the basic form of resonance curve

$$\mu \propto [(\nu-\omega)^2+\tfrac{1}{4}\gamma^2]^{-1} \tag{192}$$

where ν is the circular frequency of the radiation (rf in this case), ω the frequency of the transition, and γ is the reciprocal life time of the decaying $2p$ state. It may be recalled that in the original Weisskopf-Wigner theory, there appeared a divergent integral which had to be discarded in order to obtain Eq. (192) without an infinite shift in the resonance frequency. Serpe[86] showed that this divergent integral could be interpreted as a part of the self-energy due to the interaction of the electron with the radiation field. Presumably with the more powerful theoretical techniques[87] now available for handling divergent field theories, the problem of radiative line shape could be treated in a satisfactory way: the infinite terms would disappear and the finite level shifts of Bethe[73,75] would properly appear in the denominator. Unfortunately this does not appear to have been done as yet, but there is little doubt that a result much like Eq. (192) would be obtained for the basic resonance shape. It might be mentioned that a shift of resonance frequency such as that associated with the damping of a classical harmonic oscillator, of order $\gamma^2/\omega(2p\rightarrow 1s)$, is completely negligible for the $2p$ state of hydrogen.

60. Possible Corrections to Weisskopf-Wigner Line Form

It is still possible that (192) should be modified in some way. For instance, there might be added an antiresonant term

$$[(\nu+\omega)^2+\tfrac{1}{4}\gamma^2]^{-1}. \tag{193}$$

Indeed, such a contribution appears in some derivations, but for $\nu/2\pi\sim\omega/2\pi\sim 2400$ Mc/sec, it is 9200 times smaller than the resonant term, and is completely negligible for the present discussion.

When the derivation of Appendix II is repeated with $\mathbf{E}$ replaced by $\mathbf{E}\exp(-i\nu t)$, Eq. (76) becomes equivalent to (192). If, however, the perturbation is written in terms of a vector potential $\mathbf{A}$ of the rf field as $-(e/m)\mathbf{A}\cdot\mathbf{p}$ instead of $e\mathbf{E}\cdot\mathbf{r}$ an additional factor $(\omega/\nu)^2$ appears in (192). This would give rise to a significant distortion of the resonance curve, and it is therefore important to choose the correct form for analysis of the data. Of course, the difference between the perturbations $\mathbf{E}\cdot\mathbf{r}$ and $-(\mathbf{A}\cdot\mathbf{p})/m$ just corresponds to a gauge transformation under which the theory is known to be invariant, so that both perturbations must lead to the same physical predictions. Nevertheless, a closer examination shows that the usual interpretation[88] of probability amplitudes is valid only in the former gauge, and no additional factor $(\omega/\nu)^2$ actually occurs.

The Van Vleck-Weisskopf[89] formula for collision broadened lines does contain a factor involving ω/ν which has been confirmed[90] experimentally. There is, however, a fundamental difference between collision and radiative damping, so it should not be expected that the collision line shape would be obtained here.

61. Rf Power Shifts

Just as a shift of resonance frequency can be produced by a static electric field, it is possible for such a displacement to arise from an rf electric field, in particular, the rf field used to produce the observed quenching. A treatment of this problem is given in Appendix IV. The result is that no significant correction to the resonance position is needed for rf powers used in the precision experiments.

M. NONRADIATIVE CORRECTIONS TO LINE FORM

In Sec. 45 a number of causes for asymmetries and shifts of resonance peaks were listed. We now turn to a detailed consideration of such effects. The basic assumption is that under the influence of rf fields each metastable state decays at a rate γ given by Eq. (25). In case overlap of nearby peaks is deemed important, μ is taken as the sum of such terms. The observed beam consists of atoms with various velocities distributed among the various hyperfine components of states α and β. For the precision work, a magnetic field is chosen for which the β-contribution to the signal is at most only a few percent of that from α, so a separate correction for presence of atoms in the β-state can be made.

62. Effect of Saturation and Velocity Distribution

Of the metastable atoms in state (α, m_I) and having speed v a fraction

$$\phi = 1-\exp[-\mu(m_I, H)l/v] \tag{194}$$

is quenched while passing through an rf field of length l. This fraction must be averaged over hyperfine states and also over velocities of atoms in the beam. To a certain extent, the distribution of velocities is uncertain. Fortunately, as shown in Appendix V, the results are not appreciably dependent on this. For the calculations described below, the distribution (88) is used and gives a fractional quenching of a beam with initially un-

[85] V. F. Weisskopf and E. P. Wigner, Z. Physik **63**, 54 (1930) and **65**, 18 (1930).
[86] J. Serpe, Physica **7**, 133 (1940).
[87] See, for example, F. J. Dyson, Phys. Rev. **75**, 486 (1949).
[88] I am indebted to Professor L. H. Thomas for a helpful discussion of this point.
[89] J. H. Van Vleck and V. F. Weisskopf, Revs. Modern Phys. **17**, 227 (1945).
[90] G. Becker and S. Autler, Phys. Rev. **70**, 300 (1946).

polarized nuclear spins

$$\phi=[2U^{-4}/(2I+1)]\sum_{m_I=-I}^{I}\int_0^\infty[1-\exp\{-\mu(m_I,H)l/v\}]\times\exp(-v^2/U^2)v^3dv. \quad (195)$$

In terms of the function

$$G(\psi)=2\int_0^\infty[1-\exp-(\psi/z)]\exp(-z^2)z^3dz \quad (196)$$

which is plotted in Fig. 48 and whose properties are discussed in Appendix V, this becomes

$$\phi=[1/(2I+1)]\sum_{m_I=-I}^{I}G(\psi_m) \quad (197)$$

where

$$\psi_m=\mu(m,H)l/U. \quad (198)$$

As a starting point, we ignore overlap from other resonances and curvature of the Zeeman lines, and express the resonance parameters ψ_m in terms of frequency units ξ (megacycles per second) from their peaks

$$\psi_m=Ab^2/[(\xi-a_m)^2+b^2] \quad (199)$$

where A is proportional to the rf intensity, and $2b=\gamma$ gives the radiative width (99.692 Mc/sec for hydrogen and 99.719 Mc/sec for deuterium). The values of a_m depend on the hyperfine splitting of the initial and final states. In the approximations of Part I, for hydrogen, $a_1=-a_{-1}=58.5$ Mc/sec for αf and 29.3 Mc/sec for αe. Because of incomplete Back-Goudsmit effect, and other small corrections, the separations of the two peaks are slightly different, and improved values are used so that the spacing of the hyperfine peaks is theoretically correct. The shift of the center is taken into account subsequently together with the Zeeman curvature. For deuterium, in the above approximation $a_0=0$, $a_1=-a_{-1}=18$ Mc/sec for αf and half as much for αe.

The following procedure was adopted. All data for a given transition was taken with the rf power set to give a prescribed percentage of quenching at the center of the resonance. From the quenching at the center, it was then possible to determine the constant A of Eq. (199) independently of rf intensity, or values of $|(n|\mathbf{e}\cdot\mathbf{r}|\alpha)|^2$, l or U and to calculate the resonance curve for this value of A. When some factor leading to asymmetry or shift was to be considered, Eqs. (197)–(199) were modified accordingly, and the correction to the apparent center as measured at the prescribed level was determined. Since all considered effects led to very small shifts, it was sufficient to find such corrections independently and to add the results.

To illustrate these calculations, the case of transition αe in hydrogen at 2195 Mc/sec will be considered in detail, with results given as needed for other cases in Part IV. A panoramic view of the dependence on magnetic field of rf quenching at 2195 Mc/sec was given in Fig. 36. The peak quenching for αe was chosen arbitrarily as 31 percent, which was neither too small for accurate measurements or large enough to give serious saturation effects. As explained in Sec. 45, it was proposed to locate the center of each resonance by finding points on each side which gave equal quenching. These are called the "working points" and in the case chosen for illustration were taken at the 21 percent quenching level. It was possible to ignore the effect of overlap from the much weaker αf peak in this case.

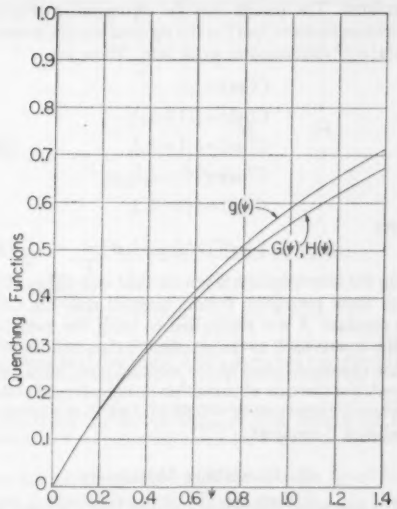


FIG. 48. Quenching functions for various assumed velocity distributions as dependent on a variable ψ which is proportional to the rate of transitions. The function $G(\psi)$ (defined in Eqs. (196) and (237)) applies for a thermal velocity distribution, while $g(\psi)$ (Eq. (241)) is obtained for an equivalent single velocity. When bombardment recoil is taken into account, there results a function $H(\psi)$ which is numerically hardly distinguishable from $G(\psi)$.

63. Variations of Matrix Element Across Resonance

The rate of decay of the metastable state α induced by rf depends on the squares of the matrix elements of $\mathbf{r}$ to the various p states. The values of these in weak magnetic field approximation are given in Table VI. At the fields used in the precision measurements, a departure from the weak field values is to be expected, and since the matrix element varies across the resonance curve because of the changing magnetic field, there is a significant amount of distortion which requires a correction.

Bethe[91] has given the intermediate field wave functions from which the matrix elements of $\mathbf{r}$ may be

[91] See reference 82, p. 398.

calculated. The results may be expressed in terms of correction factors $C(\alpha n)$ to the squared matrix elements $|(n|\mathbf{r}|\alpha)|^2$ evaluated in weak field. These are

$$\begin{aligned} C(\alpha a)&=1,\\ C(\alpha b)&=\tfrac{3}{4}(1+\delta_+),\\ C(\alpha c)&=\tfrac{3}{4}(1+\delta_-),\\ C(\alpha e)&=\tfrac{3}{4}(1-\delta_+),\\ C(\alpha f)&=\tfrac{3}{4}(1-\delta_-), \end{aligned} \tag{200}$$

where

$$\delta_\pm=(\pm\tfrac{1}{2}+x)[(9/4)\pm x+x^2]^{-\frac{1}{2}} \tag{201}$$

using the dimensionless magnetic field unit of Eq. (172). With these correction factors inserted into Eq. (199), the constant A was readjusted to bring the quenching at the center back to the standard value, and the error in the apparent center at the working point level determined. For the case of transition αe in hydrogen at 2195 Ms/sec, a correction of -0.58 Mc/sec to the apparent level shift is required.

64. Quenching Asymmetry

The preceding discussion assumed that the beam was composed of atoms divided equally among the $(2I+1)$ hyperfine states (α, m_I). If electric fields are present, these states are differently quenched because of different energy separations from competing p states. This gives rise to an asymmetry and shift of the resonance curves. The most important cause of quenching is presumably the motional electric field, and only this was taken into account in the following discussion.

The decay rate due to motional Stark effect is given as a sum of terms like (42) for transitions from α to states a, c, and f. Ignoring the slight differences in the dimensionless magnetic field variables x for the various states, the decay rate may be written as

$$\begin{aligned} \lambda=\gamma(ea_0/\mu_0)^2(V/c)^2x^2\cdot[&(9/2)(y_{a\alpha}{}^2+\tfrac{1}{4}\Gamma^2)^{-1}\\ &+(3/2)(y_{c\alpha}{}^2+\tfrac{1}{4}\Gamma^2)^{-1}+3(y_{\alpha f}{}^2+\tfrac{1}{4}\Gamma^2)^{-1}], \end{aligned} \tag{202}$$

where γ is the decay rate of $2p$, a_0 the Bohr radius, μ_0 the Bohr magneton, V the speed of the atom at right angles to the magnetic field, x is the magnetic field measured in units of 5214 gauss, while the dimensionless damping constant Γ has the value $100/7300=0.0151$. The energy separations are proportional to $y_{a\alpha}=y_\alpha=y_\alpha$, etc. and are supposed to include hyperfine splittings as well as an (approximate) value for the S level shift. In the cases of interest, however, the decay is induced mostly by state f, so that hyperfine splitting need be taken into account only in the third term. As indicated in Fig. 17, for hydrogen $y_{\alpha f}$ is increased by 58.5/7300 for $m_I=\frac{1}{2}$, and decreased by the same amount for $m_I=-\frac{1}{2}$.

Since the rate of decay due to motional electric field varies with the square of the atomic speed V, and the time spent in travelling a distance l from electron bombarder to detector varies inversely with V, only a fraction $\exp-\lambda l/V$ of the excited atoms will reach the detector in the absence of rf fields, where $\lambda l/V$ is proportional to V. The fractional quenching produced by rf may then be written as

$$\phi=\textstyle\sum_m G(\psi_m,\rho_m)/\sum_m G(\infty,\rho_m) \tag{203}$$

where

$$G(\psi,\rho)=2\int_0^\infty e^{-\rho y}(1-e^{-\psi/y})e^{-y^3}y^3dy \tag{204}$$

with ψ_m given by Eq. (198), and

$$\rho_m=\lambda^{(0)}{}_m l/U \tag{205}$$

where $\lambda^{(0)}$ is the value of λ for $V=U$, the velocity used in Eq. (88). The integral $G(\psi,\rho)$ may be evaluated by expanding $\exp-\rho y$ in a power series in ρy and integrating term by term. The first integral is just the tabulated $G(\psi)$. The subsequent integrals may be expanded in powers of ψ and integrated term by term with sufficient approximation. The values of ϕ calculated at the working points $\xi=\pm 58.8$ Mc/sec for the case at hand are 21.528 percent and 21.478 percent, respectively, and after conversion into a correction to the apparent center of the resonance curve, imply a decrease of the level shift by 0.08 Mc/sec.

65. Incomplete Back-Goudsmit Effect and Nonlinear Zeeman Splitting

At the magnetic fields used in the experiments, the nuclear spin is not fully decoupled from the other angular momenta of the atom. As a result, the hyperfine levels are unsymmetrically distributed about the energy obtained without hyperfine structure. The apparent center of the composite resonance curve is accordingly displaced from the position it would have in the absence of hyperfine structure. The correction to the level shift can be obtained by using unequal values for the a_m in Eq. (199) and calculating the apparent center. (The case of hydrogen with two component peaks could be treated more simply.)

The resonance curves are taken with fixed radio frequency by varying the magnetic field. Due to nonlinear dependence of the Zeeman splitting on magnetic field caused by progressive decoupling of **L** and **S** a distortion of the resonance arises. Letting h be the distance in gauss measured from the true center of a resonance curve, one may write with sufficient approximation in Eq. (199).

$$\xi\to\xi_m=c_mh+d_mh^2 \tag{206}$$

where c_m are the slopes of the frequency *versus* magnetic field curves for hyperfine component m, evaluated at the corresponding resonance field. In the calculations based on the ideal resonance curve this slope was evaluated at the center of the composite curve, while the curvature represented by d_mh^2 and asymmetry represented by unequal a_m were neglected. The com-

posite resonance curve was then calculated as a function of h, and the departure of the apparent center from $h=0$ was determined. For the case being illustrated, corrections of -0.81 Mc/sec and -0.11 Mc/sec, respectively, must be applied to the apparent level shift because of the incomplete Back-Goudsmit effect and the Zeeman curvature.

66. Correction for Rf Power Variation

A correction must be applied for rf power variation, both because of slow changes in oscillator output during a run, and also to bring the peak quenching exactly to the prescribed value. An approximate basis for such a correction was indicated in Eq. (102) on the assumption of a linear relationship between rf intensity and I_α and that ϕ is proportional to power. The former assumption is valid for the limited range of I_α values occurring in a run, but the second requires small ϕ. In order to test this point, $(d\phi/dA)/(d\phi/dA)_0$ was calculated for the case under study ($\phi_0=31$ percent), and found to be sufficiently near to ϕ/ϕ_0 that Eq. (102) could be used.

67. Effect of Radiation from Quenched Atoms

As described in Sec. 42, there is a background signal due to ultraviolet radiation produced in the bombardment region. Except for fluctuations the effect of this is eliminated when the ratio of rf to dc quenching is computed. In addition, there is a detector signal produced by the Lyman L_α radiation emitted in the interaction region when the metastable atoms are quenched by rf and dc fields. Fortunately this introduces no error in the fractional quenching. To see this, let η and η' be efficiencies of detection of metastable atoms and photons, respectively, and let Ω be the average solid angle subtended by the detector from the quenching region. Then if the magnitude of the unquenched beam signal is denoted by ηB, the signal received when a fraction ϕ of the metastables is quenched is

$$\eta B(1-\phi)+\eta' B\phi(\Omega/4\pi)$$

and the apparent fraction quenched is

$$\phi_{\text{app}}=\frac{\text{rf quenching}}{\text{dc quenching}}=\frac{\eta B\phi-\eta' B\phi(\Omega/4\pi)}{\eta B-\eta' B(\Omega/4\pi)}=\phi$$

so that no error results from the photons produced in the measurement of ϕ, although the dc and rf quenching are separately reduced. If one used values $\eta=0.4$ and $\eta'=0.08$ such as found by Dorrestein[92] for helium, the corrections to these would also be small since $\Omega/4\pi \sim 0.02$. If in fact η were much smaller relative to η' the apparent signal would be reduced (or even reversed in sign!). All indications are that $\eta \gg \eta'(\Omega/4\pi)$, although from the signal size it was concluded in Sec. 26 that η was probably much less than 0.4. (Observations in which large voltages of either sign are applied across the detector are compatible with the relation $\eta \sim \eta'$ but a more accurate study should be made.)

[92] R. Dorrestein, Physica **9**, 433 and 447 (1942).

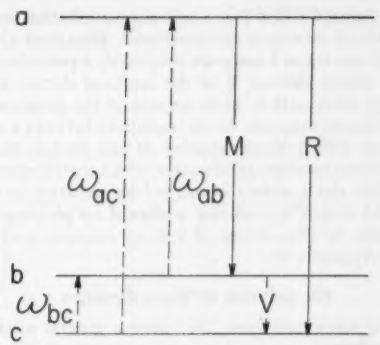


FIG. 49. Energy levels involved in sharp resonances $\alpha\beta$. Letters a and b represent states α and β (2^3S_1, $m_s=\frac{1}{2}$ and $-\frac{1}{2}$, respectively), while c represents state e (2^3P_1, $m_j=\frac{1}{2}$) which crosses β for $H=575$ gauss. The circular frequency separations ω_{ac}, etc. are shown and the perturbing matrix elements connecting various states: $\hbar R$ for electric dipole and $\hbar M$ for magnetic dipole energies due to rf fields, and $\hbar V$ the perturbing energy due to motional electric field.

N. THEORY OF $\alpha\beta$ TRANSITIONS

68. Statement of Problem

As explained in Sec. 44, it is possible to induce transitions from state α to state β which are much sharper than the transitions to nonmetastable states. These narrow resonances have been used to calibrate the magnetic field (Sec. 39), but since their appearance is rather unusual, it is necessary to have a theory of their shape in order to allow for any asymmetry shifting the apparent center.

Magnetic dipole transitions from α to β may occur because of the interaction $g_s\mu_0\mathbf{S}\cdot\mathbf{H}$ of the magnetic moment of the electron with a component of the rf magnetic field at right angles to the static magnetic field. It is also possible for electric dipole transitions to occur because the state β contains some p state contamination due to the motional electric field $\mathbf{E}=(V/c)\times\mathbf{H}$. In practice, the second mechanism is usually more important than the first.

69. Wave Equations

We consider three states $\alpha(2^3S_1,\ m_s=\frac{1}{2})$, $\beta(2^3S_1,\ m_s=-\frac{1}{2})$, and $e(2^3P_1,\ m_j=\frac{1}{2})$, as in Fig. 49, denoting their probability amplitudes by letters a, b, and c, respectively. The equations of time dependent perturbation theory are then

$$\begin{aligned} i\dot{a}&=\tfrac{1}{2}M^*b\exp[i(\omega_{ab}-\nu)t]+\tfrac{1}{2}R^*c\exp[i(\omega_{ac}-\nu)t],\\ i\dot{b}&=\tfrac{1}{2}Ma\exp[-i(\omega_{ab}-\nu)t]+V^*c\exp(i\omega_{bc}t), \qquad (207)\\ i\dot{c}&=\tfrac{1}{2}Ra\exp[-i(\omega_{ac}-\nu)t]+Vb\exp(-i\omega_{bc}t)-\tfrac{1}{2}i\gamma c. \end{aligned}$$

The circular frequency separation of state a from b is, denoted by ω_{ab} while ν is the circular radiofrequency.

As in Appendix II of Part I only the perturbation terms capable of resonance are considered. Equations (207) imply that states b and c are coupled by a perturbation with matrix element V of the motional electric field energy $e(\mathbf{V}/c)\times\mathbf{H}\cdot\mathbf{r}$, while because of the presence of radio waves magnetic dipole transitions between a and b occur with a matrix element M and electric dipole transitions between a and b occur with a matrix element R. Only the p state e is coupled by radiation to the ground state $1^2S_{\frac{1}{2}}$, and this is allowed for phenomenologically by introduction of a decay constant $\gamma=1/\tau_p$ as in Appendix II.

70. Solution of Wave Equation

The wave equations (207) have a general solution of the form

$$\begin{aligned} a&=\sum_{k=1}^{3} A_k\exp(-\mu_k t),\\ b&=\sum_{k=1}^{3} B_k\exp[i(\nu-\omega_{ab})t-\mu_k t],\\ c&=\sum_{k=1}^{3} C_k\exp[i(\nu-\omega_{ac})t-\mu_k t], \end{aligned} \tag{208}$$

where the μ_k are roots of a cubic equation

$$\begin{vmatrix} i\mu & \frac{1}{2}M^* & \frac{1}{2}R^* \\ \frac{1}{2}M & i\mu+\nu-\omega_{ab} & V^* \\ \frac{1}{2}R & V & i\mu+\nu-\omega_{ac}-\frac{1}{2}i\gamma \end{vmatrix}=0. \tag{209}$$

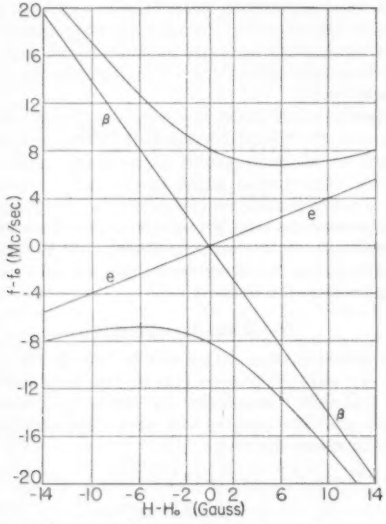


FIG. 50. Stark splitting of levels β and e as a function of magnetic field separation from the crossing field of 575 gauss. The straight lines indicate unperturbed energies and the curves show expected level splitting due to motional electric field according to usual degenerate perturbation theory.

The coefficients B_k and C_k are expressible in terms of the A_k by the equations

$$\begin{aligned} B_k&=-\tfrac{1}{2}A_k[M(\nu-\omega_{ac}+i\mu_k-\tfrac{1}{2}i\gamma)+RV^*]/\mathfrak{D}_k,\\ C_k&=-\tfrac{1}{2}A_k[R(\nu-\omega_{ab}+i\mu_k)+MV]/\mathfrak{D}_k, \end{aligned} \tag{210}$$

where

$$\mathfrak{D}_k=(\nu-\omega_{ac}+i\mu_k-\tfrac{1}{2}i\gamma)(\nu-\omega_{ab}+i\mu_k)-|V|^2.$$

The initial conditions $a=1$, $b=c=0$ at $t=0$ require

$$\sum_{k=1}^{3} A_k=1,\quad \sum_{k=1}^{3} B_k=0,\quad \sum_{k=1}^{3} C_k=0 \tag{211}$$

and these equations suffice to determine the A_k.

71. Discussion of Roots. Violation of "No-Crossing" Theorem

In practice, the roots μ_1, μ_2, μ_3 differ greatly from one another in the values of

$$\mu_k+\mu_k^*.$$

The largest of these, $\mu_1+\mu_1^*$, corresponds closely to the radiative decay rate γ of the nonmetastable state e, the intermediate value

$$\mu_2+\mu_2^*$$

is related to the Stark induced rate of decay of the lower $2^2S_{\frac{1}{2}}$ state β, while $\mu_3+\mu_3^*$ essentially determines the decay rate of the upper $2^2S_{\frac{1}{2}}$ state α caused by radio waves. If R and M are neglected, the two larger roots are solutions of the quadratic equation

$$\mathfrak{D}(\mu)\equiv\begin{vmatrix} i\mu-(\omega_{ab}-\nu) & V^* \\ V & i\mu-(\omega_{ac}-\nu+\frac{1}{2}i\gamma) \end{vmatrix}=0. \tag{212}$$

If the damping term $\frac{1}{2}i\gamma$ were absent, this would be equivalent to the secular equation for determination of the static Stark effect splitting of the approximately degenerate levels β and e. When the perturbation $|V|$ is large compared to $|\omega_{be}|$ the splitting is linear in $|V|$ while it is quadratic in $|V|$ for $|V|\ll|\omega_{be}|$. The presence of a damping term γ, however, essentially modifies the nature of the Stark effect when $\gamma\geq 4|V|$. The two roots for small $|V|$ are then approximately

$$\begin{aligned} i\mu_1&\sim\tfrac{1}{2}i\gamma+(\omega_{ac}-\nu)+|V|^2/(\omega_{be}+\tfrac{1}{2}i\gamma)\\ i\mu_2&\sim(\omega_{ab}-\nu)-|V|^2/(\omega_{be}+\tfrac{1}{2}i\gamma). \end{aligned} \tag{213}$$

The real part of the second root is given by

$$\mu_2+\mu_2^*=\gamma|V|^2/[\omega_{be}^2+\tfrac{1}{4}\gamma^2] \tag{214}$$

which is just the decay rate of the lower metastable state as calculated in Appendix II. The imaginary part

$$\tfrac{1}{2}i(\mu_2^*-\mu_2)=\omega_{ab}-\nu+\omega_{be}|V|^2/[\omega_{be}^2+\tfrac{1}{4}\gamma^2] \tag{215}$$

implies a shift in the position of that state. When $|\omega_{be}|\gg\frac{1}{2}\gamma$ the shift is just that to be expected for the quadratic Stark effect, but the shift is much reduced

for $|\omega_{be}|<\frac{1}{2}\gamma$, and in fact, even vanishes for $\omega_{be}=0$. This behavior is illustrated in Fig. 50 giving the energies of levels β and e as functions of magnetic field in the vicinity of their crossing point. In one case the electric field is zero and there is no coupling V between the states, while in the other, the usual phenomenon of energy level repulsion required by the famous von Neumann-Wigner[93] "no-crossing" theorem is shown for an electric field experienced by a 2500°K deuterium atom moving perpendicular to a magnetic field of 575 gauss. In Fig. 51 the change brought about by the radiative broadening of state e is indicated. The six solid curves correspond to six assumed values of the damping constant: (a) $\gamma=0$, (b) $\gamma=2V$, (c) $\gamma=(12)^{\frac{1}{2}}V$, (d) $\gamma=(15.36)^{\frac{1}{2}}V$, (e) $\gamma=4V\sim16$ Mc/sec (critical damping), and (f) the actual case $\gamma\sim100$ Mc/sec. The unperturbed levels are indicated by the dotted curves. It will be noted that for damping equal or above the critical value the "no-crossing" theorem is violated, and above critical damping the β-state largely retains its s character throughout. For subcritical damping, however, each level is a 50–50 mixture of s and p at the βe crossing point, and the life of each state is $2\tau_p$. The observation of sharp $\alpha\beta$ resonances at this magnetic field with a half-width of order 3 Mc/sec instead of 50 Mc/sec supports the theory of the reduced Stark splitting when one of the levels is highly damped. Such phenomena ought to occur in other problems of atomic and molecular physics, but could be observed only with instruments capable of exploring well within the radiative widths.

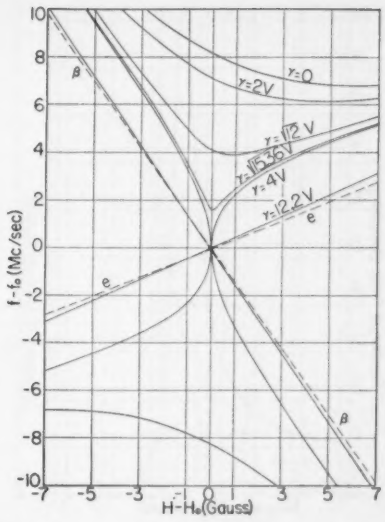


FIG. 51. Stark splitting of levels β and e as in Fig. 50 according to perturbation theory in which radiative decay of e is taken into account. Curves are shown for the motional electric field energy equal to its actual value with various assumed values for the damping constant γ. For $\gamma=0$, the curve of Fig. 50 is obtained. Critical damping occurs for $\gamma=4V$, while $\gamma=12.2V$ corresponds to the actual case with $\gamma/(2\pi)=100$ Mc/sec.

72. Calculation of Decay Constant

It will suffice, for present purposes, to calculate the small root μ_3 only to second order in R and M. One finds

$$i\mu_3=-\tfrac{1}{4}[|R|^2(\omega_{ab}-\nu)+MVR^*+M^*V^*R+|M|^2(\omega_{ae}-\nu+\tfrac{1}{2}i\gamma)]/\mathfrak{D}(0). \quad (216)$$

In the analysis of such a complicated expression, it is necessary to have a clear idea of orders of magnitude of all quantities entering the equation. The damping constant γ corresponds to a frequency of 100 Mc/sec, while the matrix element $V=(e/h)(v/c)H(e|y|\beta)$ for for deuterium at 575 gauss has a value corresponding to 8.2 Mc/sec.

If the rf field has approximately equal magnitudes of electric and magnetic fields, the matrix element R is much larger than M. In fact, $M/R\sim\mu_0/ea_0=\hbar/(2mca_0)=\frac{1}{2}\alpha\sim1/274$. Except for ν very near ω_{ab} the terms involving M in Eq. (216) may accordingly be neglected. Then

$$i\mu_3=-\tfrac{1}{4}|R|^2(\omega_{ab}-\nu)/[(\omega_{ab}-\nu)(\omega_{ae}-\nu+\tfrac{1}{2}i\gamma)-|V|^2]. \quad (217)$$

In the experiment, metastable atoms are formed in the bombardment region in both states α and β in equal numbers. After traversing a distance $L_1\sim3$ cm they enter the rf region of length $L_2\sim1$ cm and the survivors travel a further distance $L_3\sim2$ cm to the detector. For simplicity, we assume that state β has fully decayed before the atoms reach the rf region at time $t=0$. After passage through the rf region at time $t=L_2/V\sim1.2\times10^{-6}$ sec there will be a distribution of atoms in states $a=\alpha$, $b=\beta$, and $c=e$ with probabilities $|a|^2$, $|b|^2$, and $|c|^2$, respectively. The last state will strongly decay before the atoms strike the detector at a time later by $L_3/V\sim2.5\times10^{-6}$ sec. To obtain as much simplicity as possible, let us also assume that the same is true of atoms in state b. This assumption will be valid when the resonances are studied in the vicinity of the crossing point of β and e for then the decay rate $\mu_2+\mu_2^*$ is of order $\gamma/25$ as shown in Fig. 34. In the case of resonances $\alpha\beta$ studied at 2000 Mc/sec, state b is somewhat less well damped.

With these simplifications, the detector signal is proportional to

$$S\propto|A_3|^2\exp-(\mu_3+\mu_3^*)t. \quad (218)$$

The coefficient $|A_3|^2$ is less than unity (quenching of beam due to decay of transients), but is a slowly varying function of magnetic field in the vicinity of the sharp resonances so that we are more interested in the

[93] J. von Neumann and E. P. Wigner, Physik. Z. 30, 467 (1929).

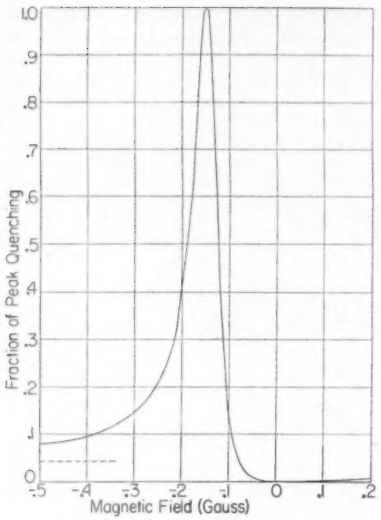


FIG. 52. Form of $\alpha\beta$ resonance expected from Eq. (219) at a frequency of 1995 Mc/sec. The fractional quenching is plotted against $H-H_0$.

rf induced decay rate given by

$$\mu_3+\mu_3^*=\tfrac{1}{4}\gamma|R|^2(\omega_{ab}-\nu)^2/ [\{(\omega_{ab}-\nu)(\omega_{ae}-\nu)-|V|^2\}^2+\tfrac{1}{4}\gamma^2(\omega_{ab}-\nu)^2]. \quad (219)$$

When $|\omega_{ab}-\nu|$ is large, and $|V|$ is sufficiently small, Eq. (219) simplifies to

$$\mu_3+\mu_3^*\rightarrow\tfrac{1}{4}\gamma|R|^2/[(\omega_{ae}-\nu)^2+\tfrac{1}{4}\gamma^2] \quad (220)$$

which is just the rate of rf induced transitions from α to e when only these two levels are present. The maximum decay rate is $|R|^2/\gamma$.

More generally, expression (219) vanishes when $\omega_{ab}-\nu=0$ and reaches a value (nearly a maximum) of $|R|^2/\gamma$ for the frequencies for which the first term in the denominator vanishes. For small $|V|$, these frequencies are near $\nu=\omega_{ae}$ for which Eq. (220) has its maximum, and $\nu=\omega_{ab}$ for which Eq. (219) has its zero.

The appearance of the resulting resonance curve (219) is quite different depending on whether $|\omega_{be}|\gg\frac{1}{2}\gamma$ or $|\omega_{be}|\ll\frac{1}{2}\gamma$. In the former case the resonance curve might appear as shown in Fig. 52. The resonance peak is displaced from the zero at $\nu=\omega_{ab}$ by an amount corresponding to the frequency $|V|^2/\omega_{be}$, just the Stark shift of state b due to interaction with c. The curve is very similar to the experimental result of Fig. 42, but the observed width is much greater than that given by Eq. (219). The explanation of this discrepancy is that the transients of Eq. (208a) are not damped highly enough to justify their complete neglect in this case. In fact, it is necessary to take the transients into account to obtain a width of resonance compatible with the uncertainty principle. Although it is not difficult to write out the more complete solution, it would still be necessary to average it over the velocity distribution in the beam, since the position of the peak depends on velocity, and this would require considerable numerical integration. Consequently the observed peak could not be readily used for a highly precise magnetic field calibration, although the error would not be large if a rough correction for Stark effect were made.

The latter case, $|\omega_{be}|\ll\frac{1}{2}\gamma$ offers more promise. In the vicinity of the crossing point $\omega_{be}=0$, the damping given by Eq. (214) is more than adequate to permit neglect of terms containing μ_3. First expectations were that the $\alpha\beta$ resonances would be very broad in this region, and subject to a large Stark effect shift of order V from the unperturbed position. As indicated in Eq. (215), however, the radiative damping of state e greatly reduces the shift and quenching of the β-state. At the crossing point of the shifts of β and e actually vanish (except for small shifts due to distant levels). Plots of Eq. (219) as a function of magnetic field in the vicinity of the crossing point are given in Fig. 53. Since there is a distribution of velocities, the shape and half-width are affected, but the zero remains at $\nu=\omega_{ab}$. The curves are highly symmetrical, easy to observe, and afford a convenient method for calibration of magnetic field in terms of frequency. An example of a curve of this type, somewhat complicated by hyperfine structure was given in Fig. 43. It will be noted that the quenching at $\nu=\omega_{ab}$ does not fall to as low a value (about two-thirds of the peak for deuterium since one-third of the beam is in a given hyperfine state) as indicated by the preceding theory. While no quantitative comparison has been attempted with a more refined theory, it is believed that this discrepancy may be attributed to neglect of the coefficient $|A_3|^2$ in Eq. (218).

The reduction in transition probability which occurs when $\nu=\omega_{ab}$ was interpreted in Sec. 47 in terms of an equivalent electrical circuit Fig. 46. A more quantum-mechanical understanding of the phenomenon can be had from an examination of the third Eq. (207) which can be satisfied when $\nu=\omega_{ab}$ for $c=0$ and $dc/dt=0$ if a and b are related by $\frac{1}{2}Ra+Vb=0$. An examination of the solution reveals that after damping of transients this relation is satisfied, and the state oscillates between a and b with such phase and amplitude relations that the decaying state c is not excited.

The author has benefited from many helpful discussions with Professor N. M. Kroll.

APPENDIX IV. RADIOFREQUENCY POWER STARK EFFECT

The possibility of a shift in resonance frequency due to the presence of rf fields was mentioned in Sec. 60. Such shifts occur in the molecular beam radiofrequency resonance method as shown by Bloch and Siegert.[94] For a spin $S=\frac{1}{2}$ undergoing Larmor

[94] F. Bloch and A. Siegert, Phys. Rev. **57**, 522 (1940), also A. F. Stevenson, Phys. Rev. **58**, 1061 (1940).

precession in a magnetic field $H_0=\hbar\omega/g_s\mu_0$ and perturbed by a perpendicular oscillating rf magnetic field $H_1\cos\nu t$, they found an increase $\delta\omega$ in the resonance frequency of

$$\delta\omega=(\omega/16)(H_1/H_0)^2=(g_s\mu_0H_1)^2/(16\hbar(\hbar\omega)). \quad (221)$$

In the present case, the induced decay of $2^2S_{\frac{1}{2}}$ is described by Eqs. (68) with V taken in the form

$$V=\hbar R\cos\nu t \quad (222)$$

and for simplicity $\gamma_a=0$ and $\gamma_b=\gamma$

$$i\dot a=R^*e^{-i\omega t}(\cos\nu t)b, \quad (223a)$$

$$i\dot b=Re^{i\omega t}(\cos\nu t)a-\tfrac{1}{2}i\gamma b. \quad (223b)$$

If damping were neglected, the problem would be exactly that solved by Bloch and Siegert, but decay changes the character of the solution considerably. In practice, the atoms spend about a microsecond in the quenching fields, or many cycles of the rf and many half-lives of $2p$. Consequently a solution of (223) in terms of a rate of induced decay of state a is needed for practical analysis of the resonance experiments.

If the nonresonant parts of $\cos\nu t=\frac{1}{2}(e^{i\nu t}+e^{-i\nu t})$ were neglected, Eqs. (223) could be solved exactly, and with suitable approximations would lead to (192). As in the Bloch-Siegert treatment, the nonresonant parts of $\cos\nu t$ are responsible for a shift of resonance frequency.

We write

$$a=e^{-\mu t}+a_1(t) \quad (224)$$

where $\mu+\mu^*$ is the desired frequency dependent decay rate of $2^2S_{\frac{1}{2}}$, and the small $a_1(t)$ allows for departures from the simple exponential law. Equation (223b) is then solved with neglect of a_1 for b subject to the initial condition $b(0)=0$ giving

$$b=-\tfrac{1}{2}R[B_1\{e^{i(\omega-\nu)t-\mu t}-e^{-\frac{1}{2}\gamma t}\}+B_2\{e^{i(\nu+\omega)t}-e^{-\frac{1}{2}\gamma t}\}] \quad (225)$$

where

$$B_1=(\omega-\nu-\tfrac{1}{2}i\gamma+i\mu)^{-1} \quad \text{and} \quad B_2=(\omega+\nu-\tfrac{1}{2}i\gamma+i\mu)^{-1}.$$

When this result is inserted in (223a), one may equate slowly and rapidly varying terms separately to zero, obtaining

$$-i\mu=-\tfrac{1}{4}|R|^2(B_1+B_2) \quad (226)$$

and

$$i\dot a_1=-\tfrac{1}{4}|R|^2\cdot[B_1e^{2i\omega t}+B_2e^{-2i\omega t}-(B_1+B_2)(e^{-\frac{1}{2}\gamma t}+e^{-\frac{1}{2}\gamma t+2i\omega t})] \quad (227)$$

whose solution subject to $a_1(0)=0$ is

$$a_1=(i/4)|R|^2[B_1(e^{2i\omega t}-1)/(2i\omega)-B_2(e^{-2i\omega t}-1)/(2i\omega)+(B_1+B_2)\{(1-e^{-\frac{1}{2}\gamma t})/(\tfrac{1}{2}\gamma)+(1-e^{-\frac{1}{2}\gamma t+2i\omega t})/(\tfrac{1}{2}\gamma-2i\omega)\}]. \quad (228)$$

The first two terms represent a rapidly oscillating contribution to a_1 of amplitude at most of order

$$B_1|R|^2/8\nu\sim|R|^2/(4\gamma\nu) \quad (229)$$

which will be seen from the subsequent numerical discussion to be negligible for the rf power used in the precision experiments. The remaining terms vanish at $t=0$, but after $t=1/\gamma$ very rapidly take on a constant value. Besides the "secular" damping $1-\exp-(\mu+\mu^*)t$, there is according a "shock" damping which for $(\mu+\mu^*)t\ll1\ll\gamma t$ is approximately

$$-(a_1+a_1^*)\to\tfrac{1}{2}|R|^2(B_1+B_1^*)/\gamma\simeq\tfrac{1}{2}|R|^2/[(\omega-\nu)^2+\tfrac{1}{4}\gamma^2] \quad (230)$$

and is small compared to the secular term by a factor $2/(\gamma t)$ or about 1/400 in practice, and in any case leads to no asymmetry. It can also be easily verified that no significant error was made by neglecting a_1 in Eq. (223b) as long as

$$|R|^2/\gamma^2\ll e^{-\mu t}. \quad (231)$$

In order to evaluate the damping constant μ we set $\mu=\mu_r/2+i\mu_i$ and find with sufficient accuracy

$$\mu_i=-\tfrac{1}{4}|R|^2[(\nu-\omega)/\{(\nu-\omega)^2+\tfrac{1}{4}\gamma^2\}-(\nu+\omega)/\{(\nu+\omega)^2+\tfrac{1}{4}\gamma^2\}], \quad (232)$$

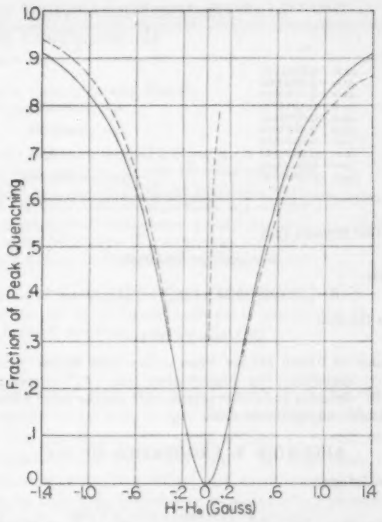


FIG. 53. Form of $\alpha\beta$ resonance expected from Eq. (219) at frequencies of 1610 Mc/sec (at βe crossing) (solid curve) and 1615 Mc/sec (5 Mc/sec away from βe crossing) (dotted curve). The latter curve is slightly asymmetrical and the apparent center as a function of quenching level is shown.

$$\mu_r=\tfrac{1}{4}\gamma|R|^2[\{(\nu-\omega-\mu_i)^2+\tfrac{1}{4}(\gamma-\mu_r)^2\}^{-1}+\{(\nu+\omega)^2+\tfrac{1}{4}\gamma^2\}^{-1}]. \quad (233)$$

The rate of decay is given by μ_r. As indicated in Sec. 60, the antiresonant contribution is negligible. The peak position is determined by

$$\nu=\omega+\mu_i\simeq\omega-|R|^2(\nu-\omega)/\gamma^2+|R|^2/8\omega, \quad (234)$$

where ν has been replaced by ω in the denominators of (232). The second term of (234) produces no shift in peak position since it vanishes for $\nu=\omega$, but it does lead to a slight reduction[96] in half-width of relative order $|R|^2/\gamma^2$ as does the term with μ_r in the denominator of (233). The last term

$$|R|^2/8\omega$$

gives an rf quadratic Stark effect shift similar in form to that for static electric fields except for a factor $\frac{1}{2}$ which can perhaps be understood from the circumstances that the antiresonant perturbation has an amplitude $(1/2)R$ and the frequency denominator is effectively 2ω.

When the shift

$$|R|^2/8\omega$$

is written in the notation used by Bloch and Siegert, it is found to be one-half as large as theirs. This difference may be attributed to the damping of one of the states which in fact is responsible for the relative simplicity of our derivation. It can be seen from Eq. (228) that $a_1(t)$ would not be negligible if there were no damping.

The numerical value of the rf Stark shift $|R|^2/8\omega$ and the dimensionless ratio $|R|^2/\gamma^2$ will now be estimated for an rf power sufficient to give 63 percent quenching at resonance in 1.25×10^{-6}

[96] This "power sharpening" is at least in part due to mixing of the long-lived $2S$ state with the decaying $2P$ state giving a reduction of the radiative width of $2P$.

TABLE VII. Coefficients of series expansion for $G(\psi)$.

c_n	b_n
$c_1 = 0.8862269$	
$c_2 = -0.5000000$	
$c_3 = 0.2954090$	
$c_4 = -0.1014592$	$b_4 = 0.0825000$
$c_5 = -0.0295409$	
$c_6 = 0.0057894$	$b_6 = -0.0027778$
$c_7 = 0.0004689$	
$c_8 = -0.0000645$	$b_8 = 0.0000248$
$c_9 = -0.0000026$	
$c_{10} = 0.0000003$	$b_{10} = -0.00000001$

sec. This requires that

$$\mu_r = |R|^2/\gamma = 8\times 10^5 \text{ sec}^{-1}$$

so that

$$R = [(8\times 10^6)(6.25\times 10^9)]^{\frac{1}{2}} = 2.24\times 10^7 \text{ sec}^{-1}. \quad (235)$$

Then the shift

$$|R|^2/8\omega \sim 1\times 10^4 \text{ sec}^{-1} \quad (236)$$

amounts to 0.0016 Mc/sec when $\omega/2\pi = 1000$ Mc/sec which is entirely negligible. The dimensionless ratio $|R|^2/\gamma^2$ giving the "shock" damping and power sharpening[95] has the value 1.3×10^{-2} and causes no significant error.

APPENDIX V. PROPERTIES OF $G(\psi)$

The integral

$$G(\psi) = 2\int_0^\infty [1-\exp(-\psi/z)]e^{-z^2}z^2dz \quad (237)$$

of Sec. 62 obeys the differential equation

$$(\psi/2)G'''-G''+G-1=0 \quad (238)$$

with boundary conditions at the regular singular point $\psi=0$, $G(0)=0$, $G'(0)=(1/2)(\pi)^{\frac{1}{2}}$, $G''(0)=-1/2$, and $G'''(0)=(1/6)(\pi)^{\frac{1}{2}}$, while $G(\infty)=1$. The fourth and higher derivatives of G are infinite at $\psi=0$, and one finds an expansion about that point in the form

$$G(\psi) = \sum_{n=1}^{\infty} c_n\psi^n + (\log\psi)\sum_{n=4}^{\infty} b_n\psi^n \quad (239)$$

with coefficients as given in Table VII.

The series converges rapidly for the values of ψ needed in the experiment. For very large ψ the saddle point approximation to (237) may be used

$$G(\psi) \sim 1-(\pi/3)^{\frac{1}{2}}\psi\exp[-3(\psi/2)^{\frac{1}{3}}]. \quad (240)$$

The relative unimportance of the exact form of velocity distribution is shown in Fig. 48 where $G(\psi)$ is compared to the functions resulting from two other plausible assumptions. The first

$$g(\psi) = 1-\exp[-\tfrac{1}{2}\pi^{\frac{1}{2}}\psi] \quad (241)$$

would describe the quenching if all atoms had a velocity $V=\frac{1}{2}\pi^{\frac{1}{2}}U$ equal to the reciprocal average reciprocal velocity. The second function $H(\psi)$ shows the quenching if the recoil at threshold for hydrogen is taken into account.

The differences are not great, but the functions had to be calculated to be sure of this. In addition, the related function $G'''(\psi)$ plays a more significant role in discussion of similar experiments on singly ionized helium.

APPENDIX VI. MISCELLANEOUS ENERGY LEVEL CORRECTIONS

(*a*) *Finite Size of Deuteron.* The finite size of the deuteron leads to an elevation of $2S$ relative to $2P$ by an amount

$$(1/48)\alpha^4 hcR(h^3/Mr_0^2I), \quad (242)$$

where I is the binding energy of the deuteron and $r_0=\alpha^2 a_0$ is the classical electron radius. For simplicity, a zero range wave function has been used for the deuteron. Numerically, this shift amounts to $+0.45$ Mc/sec, and will be contained in the experimentally determined value of $\mathcal{S}$ for deuterium. The same will be true of the other corrections of order $\mathcal{S}/M$ such as that mentioned in reference 80.

Since the deuteron has a quadrupole moment, $Q=2.73\times 10^{-27}$ cm^2 there are also contributions to the energy of $2^2P_{\frac{1}{2}}$, of an amount

$$(1/40)(Q/r_0^3)\alpha^4 hcR[m_J^2-(5/4)][m_I^2-(2/3)]. \quad (243)$$

At most this energy amounts to 0.006 Mc/sec and may be neglected.

(*b*) *Meson Cloud about Nucleus.* Slotnick and Heitler[96] have shown that the distribution of mesons about a proton gives a $2S$ level shift of about 0.02 Mc/sec according to a typical meson theory consistent with observations on the electron-neutron interaction.

(*c*) *Correction to Hyperfine Structure Splitting.* The hyperfine structure for $n=2$ has been expressed in Sec. 56 as a multiple of Δw, the splitting for $2^2S_{\frac{1}{2}}$. This in turn was taken to be one-eighth of the hyperfine splitting for $1^2S_{\frac{1}{2}}$ as measured by Prodell and Kusch.[63] Any correction due to the difference in binding ought to be far below the level of accuracy needed here. Another neglect involved the replacement of the coefficient 2 of $\mathbf{I}\cdot\mathbf{L}$ in Eq. (177) by g_S. This implies a correction to Eq. (177) of

$$(J|w_{11}'|J) = -(g_S-2)[\Delta w/(I+\tfrac{1}{2})]m_I m_J \cdot \begin{cases} 1/12 & 2^2P_{\frac{3}{2}} \\ 1/24 & 2^2P_{\frac{1}{2}} \end{cases} \quad (244)$$

but it amounts at most to 0.013 Mc/sec, and has been neglected.

[96] M. Slotnick and W. Heitler, Phys. Rev. **75**, 1645 (1949).

Dissociation of Hydrogen Molecules by Vibrational Excitation and Three-Body Recombination Coefficient. II

E. BAUER*
Division of Physics, National Research Council of Canada, Ottawa, Canada
(Received October 8, 1951)

The recombination of two hydrogen atoms in the presence of a third in a thermal collision is considered, the principle of detailed balance being used to obtain the three-body coefficient C_3 from the cross section σ_2 of the reverse, dissociation process. This cross section is calculated by a method developed previously which depends on correcting the Born approximation by means of the exact one-dimensional problem. We consider the case when the molecule is formed in a highly excited vibrational state of the electronic ground state and the excess energy of several kT goes into translational energy of the third atom. The three-body coefficient in this case is of order 3×10^{-36} cm^6 sec^{-1}, which is still much smaller than the empirical values. Some estimate is made of the validity of the correction method.

The excitation and de-excitation of the first vibrational state of hydrogen by collisions with protons is investigated, and it is found that an excited molecule requires 10^2 to 10^4 thermal collisions to lose its excitation; this is roughly of the correct order of magnitude for such processes.

I. INTRODUCTION

CHEMICAL recombination reactions typified by the process

$$H+H+X\rightarrow H_2+X^*, \quad (1)$$

where H is a hydrogen atom, and X is an atom or molecule, are of importance in the upper atmosphere and perhaps also in chemical kinetics, but have not been studied extensively theoretically. There are three basic kinds of mechanism, namely:

(a) The two hydrogen atoms come together in the repulsive $^3\Sigma$ electronic state of the molecule and give up their excess electronic energy to X.

(b) They may come together in the vibrational continuum of the normal $^1\Sigma$ electronic state and transfer their excess vibrational energy to X.

(c) They may come together with a large orbital angular momentum, and transfer some of this rotational energy to X.

In each case different mechanisms are possible depending on the mode of excitation of X. In case (a), the most probable process is one involving the electronic excitation of X, which has been studied recently by Moses and Wu.[1] In thermal energy collisions electronic processes are adiabatic, and in case (b) the transfer of vibrational energy to the translational excitation of X has been studied by the present writer.[2] The probability of the dissociation of H_2 from the vibrational ground state by excitation into the vibrational continuum of the normal electronic state in collision with a hydrogen atom (Section VI) and a proton (Section II) was calculated, and the probability of (1) estimated from this by using the principle of detailed balance; the resulting three-body recombination coefficient was found to be of the order 10^{-43} cm^6 sec^{-1}.

Experiments[3] give the value $C_3\sim3\times10^{-32}$ cm^6 sec^{-1}. This seems to be very large, since even a gas kinetic value for the dissociation cross section σ_2 would only correspond to $C_3\sim10^{-35}$ cm^6 sec^{-1}. In any case, the difference between theoretical and experimental values is so large that it is desirable to examine other processes theoretically. A process similar to the one considered previously,[2] in which the final state of the molecule is not the vibrational ground state but rather a highly excited vibrational state, may be much more probable as the energy to be transferred is then much smaller. This is supported by observations that in the afterglow in nitrogen in which there is no electrical excitation and the bands presumably arise from molecules resulting from recombinations, bands from highly excited vibrational levels are observed. In the present work, an estimate has been made of the probability of the transfer of an amount of vibrational energy of the order of the thermal energy in collisions of type (1), where X is a proton or hydrogen atom. The method of calculation is that described previously.[2]

The cross section for excitation of the first vibrational state from the ground state is calculated, and also that for the $v=13$ to $v=14$ excitation. The first of these gives information on the probability that a thermal collision of a proton with a vibrating molecule de-excites the vibration, while the second cross section would be expected to be of the same order as the cross section σ_2 of the reverse process of (1), as the energy transfer and other conditions are rather similar.

II. THREE-BODY COEFFICIENT

We consider transitions between a bound state Q of the molecule of binding energy Δ_Q, and a free state P in which the two atoms fly apart with relative kinetic energy $E_p=\hbar^2p^2/2M_0(M_0=\frac{1}{2}M_H)$; the third body (proton) has initially momentum $\hbar\mathbf{k}_0$, finally $\hbar\mathbf{k}_f$, all in the center-of-mass system in which its reduced

* National Research Laboratories Postdoctoral Fellow. Contribution No. 2624 from the National Research Council, Ottawa, Canada.

[1] H. E. Moses and T. Y. Wu, Phys. Rev. **83**, 109 (1951), Appendix 5.

[2] E. Bauer, Phys. Rev. **84**, 315 (1951).

[3] H. J. Schumacher, *Chemische Gasreaktionen* (Theodor Steinkopff Verlag, Leipzig, 1938), p. 324.

mass is $M_A=\frac{2}{3}M_H$. The interaction between the molecule in the $P\to Q$ transition and the atom is taken to be[4,2]

$$V_{QP}(r)=Ze^2\frac{\eta_{QP}(0.42)}{\rho_0}\Lambda_\mu\frac{e^{-\mu r}}{r}, \tag{2}$$

where $\Lambda_\mu=(1-\frac{1}{2}\mu\partial/\partial\mu)$, ρ=nuclear separation$=\rho_0+\eta$, ρ_0=equilibrium nuclear separation$\simeq$0.75A, $Z=1.115$, and $\mu=2.332$.

The cross section σ_2 for the $(Q,0)\to(P,f)$ transition is obtained by using the Born approximation and correcting it by multiplying by ζ^2, the square of the ratio of exact to sine wave matrix element of the interaction in one dimension, as once again the Faxén-Holtsmark method is not practicable because of the large number of phases required. The resulting expression for σ_2 is[2]

$$\sigma_2(\Delta_Q,k_0)=\int dE_p\cdot\eta_{QP}^2(0.42/\rho_0)^2\chi(k_0,k_f(p),p)\zeta^2,$$

$$\chi=\pi\alpha_0^2(2/\mu)^2A,\quad A=[Y+Y^2+\tfrac{1}{3}Y^3]Y_+^-, \tag{3}$$

$$Y_\pm=1/(1+X_\pm),\quad X_\pm=[(k_0\pm k_f)/\mu]^2,$$

$$\alpha_0=Ze^2/\hbar v_0=ZM_Ae^2/\hbar^2k_0.$$

χ is a factor determined by the contribution of the translational motion to the cross section, evaluated by the use of Born's approximation. The state P is in the continuum, and is represented by a wave function normalized per unit energy range, so that integration over E_f is necessary. The range of integration is determined by energy conservation—

$$\hbar^2/2M_A(k_0^2-k_f^2)=\Delta_Q+(\hbar^2p^2/2M_0), \tag{4}$$

and goes from $p=0$ to the value corresponding to $k_f=0$.

The correction factor ζ is evaluated as before[2] using as "exact" wave functions Bessel functions of order chosen to represent the correct classical turning point of the motion, X_1. This should be a good approximation, since most of the contribution to the matrix element comes from the region near X_1. Some values of ζ are given in Table I, and some discussion of the method is given in Sec. IV.

To evaluate η_{QP}, we suppose that the internuclear potential (a function of ρ) may be represented by a square well of depth D (the dissociation energy) and width l_1. Then we define

$$\begin{aligned}\hbar^2Q^2/2M_0&=D-\Delta_Q,\\ \hbar^2P^2/2M_0&=D+\hbar^2p^2/2M_0.\end{aligned} \tag{5}$$

In order to have a bound state Q, it is necessary that

$$Ql_1=(n+\tfrac{1}{2})\pi,\quad n\text{ an integer}, \tag{6}$$

and taking $\Delta_Q=kT$ (at 300°K), $l_1=1.91a_0$, we have $Ql_1=10.5\pi$, and the wave functions are

$$\begin{aligned}\psi_Q(\rho)&=N_Q\cos[Q(\rho-\rho_1)-\pi/4],\\ N_Q&=[2/(l_1+1/Q)]^{\frac{1}{2}}\simeq(2/l_1)^{\frac{1}{2}},\\ \psi_P(\rho)&=N_P\cos[P(\rho-\rho_3)-\pi/4],\\ N_P&=(1/\hbar)(2M_0/\pi P)^{\frac{1}{2}}.\end{aligned} \tag{7}$$

ψ_Q is normalized to unity within the well, which goes from $\rho_1\lesssim\rho\lesssim\rho_1+l_1$, while ψ_P is normalized per unit energy range. The evaluation of η_{QP}^2 is straightforward; we are only interested in values $E_p\lesssim kT$, and there η_{QP}^2 has practically the same value as for $p=0$, namely, if we put $\rho_1\simeq\rho_0$,[4a]

$$\eta_{QP}^2=M_0l_1^4/4\pi\hbar^2Ql_1. \tag{8}$$

In evaluating σ_2 it is convenient to change the integration over E_p into one over $E_f=\hbar^2k_f^2/2M_A$, and the result is

$$\sigma_2(\Delta_Q=kT;\ E_0'=4.5kT)\simeq6\cdot10^{-19}\text{ cm}^2.$$

E_0' is given in the "laboratory" system of coordinates in which the center of mass of the molecule is at rest, rather than in the relative system of coordinate in which the calculation is carried out; thus $E_0'=\frac{3}{2}E_0$.

Of more interest than σ_2 is the three-body recombination coefficient $C_3(E_p,E_f)$ which gives the probability that two hydrogen atoms moving with relative kinetic energy E_p will recombine on colliding with a proton whose kinetic energy is E_f relative to the center of mass of the system. C_3 may be found in terms of σ_2 by means of the principle of detailed balancing; if we write

$$\sigma_2(\Delta_Q,k_0)=\int_0^{p_{\max}}pdp\Sigma(k_0,k_f(p),p), \tag{3'}$$

then, neglecting angular factors,

$$C_3(E_p,E_f)=\pi\hbar\frac{G_{H_2}}{G_H^2}\frac{G_{X^*}}{G_X}\frac{M_0\rho_0^2}{M_A^2}\frac{1}{pk_f}\int_0^{k_2}\times\Sigma(k_0,k_f(p),p)k_0^2dk_0, \tag{10}$$

where k_2 is determined by conservation of energy [Eq. (4)]. The G's are statistical weights, $G_H=2$, $G_{H_2}=1$, $G_X=G_{X^*}$. If $E_p=kT$ and E_f', the kinetic energy of the proton relative to the molecule, is $\frac{3}{2}kT$, we find $C_3\simeq3\times10^{-37}$ cm^6 sec^{-1}.

III. THE EXCITATION OF SOME DISCRETE VIBRATIONAL STATES

It is of some interest to calculate the cross section for transitions between two discrete vibrational states a, b of the electronic ground state in slow collisions with a proton. For a transition $(k_0,a)\to(k_f,b)$,

$$\sigma_{ab}=\pi\alpha_0^2|\eta_{ab}|^2(2/\mu)^2A\zeta^2(0.42/\rho_0)^2, \tag{11}$$

[4] T. Y. Wu, Phys. Rev. **71**, 111 (1947).

[4a] This approximation is very rough, but if one describes the inter-nuclear potential by a Morse function one gets approximately the same result.

TABLE I. The correction method for protons.

	E_f'(xv)	α_0	α_f	X_1^2/a_0	X_1^2/a_0	n_0	n_f	ζ^2	σ(cm^2)	σ/σ_{max}
A	Large	1/137	1/137	1	1	½	½	1		1
B	165 ev	13.7	14	0.212	0.218	21.2	21.5	0.152	3×10^{-25}	5
B	41.5 ev	27.5	30	0.522	0.572	25.8	26.2	0.035	2×10^{-25}	0.4
C	1.65 ev	137	192	1.66	1.93	16.6	13.8	0.0075	10^{-19}	0.04
C	1.05 ev	171	354	1.84	2.45	14.7	9.5	0.0039	2×10^{-20}	0.005
D	6 kt	450	900	2.65	3.20	8.1	4.9	0.0035	8×10^{-19}	0.06
E	3.9 kt	560	1200	2.81	3.44	6.9	3.95	0.0026	3×10^{-19}	0.03
F	3.9 kt	560	560	2.81	2.81	6.9	6.9	0.0026	3×10^{-17}	0.05

* A = Very high energies, B = Dissociation from the ground state; energy transfer 4.5 ev (reference 2); $\sigma=\sigma_{12}$. C = Excitation of the first vibrational state; energy transfer 0.5 ev (Sec. III); $\sigma=\sigma_{01}$. D = Transition $v=13$ to 14; energy transfer 621 cm$^{-1}\simeq 3kT$ (Sec. III); $\sigma=\sigma_{13\to14}$. E = Dissociation from a highly excited vibrational state; energy transfer $2kT$ (Sec. II); $\sigma=\sigma_{12}$. F = Elastic collision, quoted for comparison with E. E_f' = kinetic energy of the proton relative to the center of mass of the molecule. α_{00}, α_f = coupling constants $Ze^2/\hbar v$ for initial and final states. x_1^2/a_0, x_1^2/a_0 = classical closest distance of approach for initial and final states measured in units of the Bohr radius $a_0=\hbar^2/me^2\simeq0.529\times10^{-8}$ cm. n_0, n_1 = order of the Bessel function in the "exact" one-dimensional evaluation. ζ = correction factor: ratio of "exact" to sine-wave matrix element of the interaction (2) in one dimension. σ = cross section. σ_{max} = maximum value, defined by Eq. (13).

and detailed balancing gives

$$\sigma_{ba}=(E_0/E_f)\sigma_{ab}. \tag{12}$$

The $v=0\to v=1$ transition is of some practical interest, as the probability of de-excitation of the first vibrational state of molecules may be determined experimentally from measurements on the dispersion of high frequency sound waves in gases.[5] At thermal energies $\sigma_{10}\simeq10^{-19}$ cm^2, so that one in 10^3 to 10^4 collisions with a proton should de-excite the vibration. The vibrational dispersion has not been observed in hydrogen where the vibrational frequency is relatively large, but for other gases generally one in 10^3 to 10^6 collisions with gas molecules produces de-excitation.

The cross section for the transition between the two highest observed vibrational levels, $v=13$ and 14[6] has also been worked out. Here the energy difference is 621 cm$^{-1}\simeq3kT$. Instead of calculating $\eta_{vv'}$ one works out $\langle v|d/d\eta|v'\rangle$, using wave functions derived from a Morse potential that fits the dissociation energy and vibration frequency for low excited states, as this has been found much easier than to evaluate $\eta_{vv'}$ directly.[7] From the wave equation one obtains readily the relation

$$\eta_{vv'}=\frac{\hbar^2}{M_0(E_v-E_{v'})}\langle v|d/d\eta|v'\rangle,$$

and it is found that

$$\sigma_{13\to14}(E_f'=\tfrac{3}{2}kT)\simeq8\times10^{-18}\text{ cm}^2.$$

The large value as compared with σ_{01} arises physically from the relatively small energy transfer in the collision. $\sigma_{13\to14}$ is about ten times as big as σ_2 of the previous section; the difference probably comes mostly from the different vibrational wave functions used in the two cases. If one uses $\sigma_{13\to14}$ instead of σ_2 in the relation of detailed balance (10), one obtains $C_3\simeq3\times10^{-36}$ cm^6 sec^{-1}, which is still much smaller than the experimental values.

[5] W. T. Richards, Revs. Modern Phys. **11**, 33 (1939).

[6] H. Beutler, Z. Phys. Chemie, **B27**, 287 (1934).

[7] T. Y. Wu, *Vibrational Spectra and Structure of Polyatomic Molecules* (Edward Brothers, Ann Arbor, Michigan, 1946), second edition, p. 73. Strictly speaking, the use of a Morse potential is not justified for these highly excited states; nevertheless, one would expect to get the correct order of magnitude for $\eta_{vv'}$.

IV. MAXIMUM VALUES FOR CROSS SECTIONS

It is well known that the conservation of particle flux sets maximum values to the partial cross sections of given orders.[8] In the present case the use of Wu's argument[4] gives

$$\sigma_{ab}\leqslant4\pi R^2(0.42\eta_{ab}/\rho_0)^2. \tag{13}$$

For low energy collisions of protons the range R of the interaction may be taken as the classical closest distance of approach, X_1. Some values of σ/σ_{max} are given in Table I.

V. DISCUSSION OF THE CORRECTION METHOD: PROTON AS THIRD BODY

It is not possible to say anything very definite about the method of correcting the Born approximation by multiplying the cross section by ζ^2, where ζ is the ratio of exact to sine-wave matrix elements of the interaction in one dimension, but several remarks may be made.

(a) For coupling constants $\alpha<1$ we find $\zeta\sim1$, so that the method is satisfactory for sufficiently high energies.

(b) At low energies the correction will become better as fewer phases are required; in fact the method should be satisfactory if only the S-wave had to be considered. (However, the reason for using the method at all is that in all practical cases a large number of phases is required.)

(c) As a test of the method, the calculated cross sections are compared with the maximum cross sections of the previous section in Table I. At energies less than 80 ev all calculated cross sections lie well within the maximum permissible values, while the uncorrected Born approximation would be far in excess of the limit.[9]

(d) ζ is not a very critical function of the different parameters.

[8] N. F. Mott and H. S. W. Massey, *Theory of Atomic Collisions* (Clarendon Press, Oxford, 1949), Chapter 8.

[9] For $1<\alpha_0<20$ the cross section calculated by this method exceeds the maximum by a factor up to 15. This particular energy range is not very interesting, corresponding to 30,000 ev$>E_0'>$80 ev. We may estimate that the correction leads to an uncertainty of perhaps a factor 100 in the cross section.

VI. THE CASE OF HYDROGEN ATOMS AS THIRD BODY

To the present approximation,[2] the difference between protons and hydrogen atoms is that the orbital electron produces an (attractive) screening field; electronic transitions are of course not considered here. The effect of the screening field is to make the total interaction with a hydrogen molecule weakly attractive for separations greater than $1.3a_0$, with a maximum depth of $-0.02e^2/a_0$ at $r=1.8a_0$; at smaller distances ($r\lesssim a_0$) the effect of the screening field is negligible, and the field is again repulsive.

In the case considered previously[2] the attractive potential is of no importance at the lower energies ($\sim$5 ev), although it does give a contribution at large energies. In the present case the closest distance of approach for protons is greater than $1.3a_0$, so that the screening field affects the one-dimensional matrix element very strongly. If the correction method is applied in the usual way, one finds a correction factor $\zeta\sim(+)10^{-3}$ to 10^{-2} for the (0→1) transition, i.e., for energies $\gtrsim$1 ev, and $\zeta\sim(-)10^{-2}$ for thermal energies (dissociation and elastic scattering). The change of sign of ζ arises from the change of sign of the potential with increasing distance. The calculation is relatively difficult, because it is impossible to use the Bessel function method and the much more tedious WKB method must be used instead. The result is that for thermal energies, elastic scattering cross sections and three-body coefficients are increased by a factor of ten over those for protons, while the (0→1) cross sections are not changed in their order of magnitude, although the change of sign of ζ with the corresponding zero in the cross section falls not far above the threshold of the (0→1) transition.

The conclusion is that we may perhaps accept the thermal energy cross sections very tentatively: the three-body coefficient for two hydrogen atoms recombining in the presence of a third is of order 3×10^{-36} cm^6 sec^{-1}, while the cross section for the elastic scattering of a hydrogen atom of energy $\frac{3}{2}kT$ by a hydrogen molecule is of order 5×10^{-15} cm^2. Thus we see that it is still not possible to understand the experimental value of the three-body coefficient; electronic resonance processes are hardly likely as most other gases (nitrogen, halogens) give similar values for C_3 as does hydrogen.[3]

I should like to thank Dr. T. Y. Wu for suggesting this problem and for many helpful suggestions.

PHYSICAL REVIEW VOLUME 85, NUMBER 2 JANUARY 15, 1952

An Investigation of Alpha-Particle Groups from $Al^{27}(d,\alpha)Mg^{25}$*

E. C. Toops, M. B. Sampson, and F. E. Steigert
Department of Physics, Indiana University, Bloomington, Indiana

(Received July 30, 1951)

The alpha-particle groups produced by the bombardment of aluminum by 10.8-Mev deuterons have been measured using nuclear emulsion plates for detection. Seven new levels in Mg^{25} were found at 6.98, 7.85, 8.62, 9.06, 975, 10.78, and 11.89 Mev. The previously known level at 4.01 Mev has been resolved into two levels at 3.96 and 4.12 Mev.

INTRODUCTION

SEVERAL investigations have been made of the lower-lying levels in Mg^{25} from the $Al^{27}(d,\alpha)Mg^{25}$ reaction.[1] Recently Schelberg, Sampson, and Cochran[2] extended this work to show ten levels in Mg^{25} with a maximum excitation of 6 Mev.

The present investigation was initiated with the aim of extending this work to higher excitations since the previous results showed an approximately constant level spacing of about 0.6 Mev. The energy levels above 7 Mev in Mg^{25} have been obtained in neutron scattering experiments,[3] and it seemed of interest to extend the $Al^{27}(d,\alpha)Mg^{25}$ work to overlap this region.

The present investigation was carried out with the use of nuclear emulsions as detectors to facilitate the measurement of lower energy alpha-particles leading to higher excitation in the residual nucleus. The results for lower energy alpha-particles were limited however, by the presence of background tracks.

EXPERIMENTAL PROCEDURE

The deuteron bombardments of the target were made in the same way as previously described.[2] The deuterons from the Indiana University cyclotron were passed through the analyzing magnet before striking the target foil. The beam energy was thus determined to be 10.82 ± 0.08 Mev and was held at this value for the duration of the bombardment.

The targets were of aluminum foil, 0.17 mg/cm^2 surface density, and were supported by a brass ring in such a manner that the beam was prevented from striking the brass. Fresh targets were prepared for each bombardment to reduce the accumulation of contami-

* This work was assisted by the joint program of the ONR and AEC.

[1] E. McMillan and E. O. Lawrence, Phys. Rev. **47**, 343 (1935); Pollard, Sailor, and Wyly, Phys. Rev. **75**, 725 (1949); A. P. French and P. B. Treacy, Proc. Phys. Soc. (London) **63**, 665 (1950).

[2] Schelberg, Sampson, and Cochran, Phys. Rev. **80**, 574 (1950).

[3] D. E. Alburger and E. M. Hafner, Revs. Modern Phys. **22**, 373 (1950).

nants such as pump oil. The target holder was arranged to rotate about its central axis. It was set so the plane of the foil made equal angles with the beam and the camera.

The reaction particles resulting from the bombardment were detected by the use of 1×3 inch, 50-micron nuclear emulsion plates (Kodak NTA No. 443,959–24). The plates were exposed in a conventional type of camera. A removable brass plate holder was designed to press down on the edges of the emulsion, which prevented peeling of the plates during exposure in vacuo. Two adjustable slits separated by 2.5 cm provided collimation of the particles entering the camera. The slit opening used in the present work was such as to give an angular resolution of plus or minus one-half degree. The plate holder was held in the camera so that the surface of the emulsion made an angle of 7° with the axis of the collimating slits.

The camera was designed to be used in the bombardment chamber previously used,[2] and could be set at any angle with an accuracy of 15 minutes between 15° and 140° in the laboratory coordinates. The incident beam from the analyzer was defined by a system of slits to limit the energy spread to about ±40 kev for the beam energy of 10.8 Mev.

A number of plates were exposed at various angles for the different lengths of time. The total exposure was determined by monitoring the beam current to the target with a current integrating circuit. The exposures ranged from 3 to 14 microcoulombs of charge.

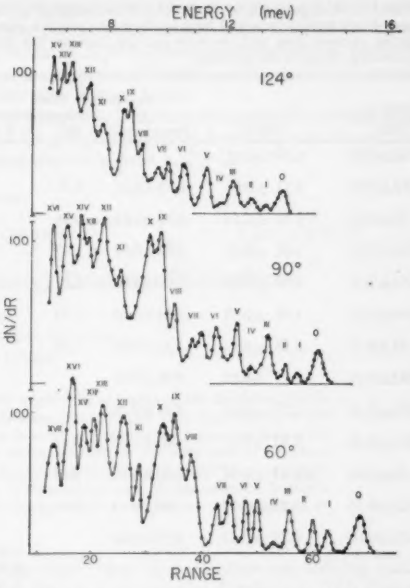


FIG. 1. Plots of the range of alpha-particle tracks in the emulsion vs the number per range interval for three plates taken at 60°, 90°, and 120° (1 unit = 1.674 microns).

PROCESSING AND MEASUREMENT OF PLATES

Standard processing procedures were followed, developing the plates for six minutes in full strength D-19b developer at 18°C, and fixing for one hour. It was found that the use of a hardening bath following development (10 min in Kodak SB-3) prevented the tendency of the emulsion to form blisters.

The alpha-particle track lengths were measured with a Spencer 5LP binocular research microscope. The left eyepiece contained an eyepiece scale which was calibrated against a stage micrometer. The right body tube was fitted with a simple angular scale, which enabled tracks with excessive angular divergence to be rejected.

The presence of large numbers of protons, from competing reactions, and scattered deuterons made the measurement of alpha-particle tracks in the emulsion difficult. An attempt was made to utilize underdevelopment of the plates to give better discrimination between the different types of tracks. This method proved unsatisfactory as the visibility of the alpha-particle tracks became so low as to make the observations very tedious and slow.

The phenomenon of fading of the latent image of tracks in nuclear emulsions on delayed development[4] provides one with a method of eliminating the unwanted proton and deuteron tracks, since the fading is much more marked for these tracks than for alpha-particle tracks.

The procedure used on most of the plates was to delay development for one day. This was sufficient to nearly eradicate the proton tracks and still allow the alpha-particle tracks to remain. The preponderance of scattered deuterons in the forward direction makes the use of nuclear emulsion detection particularly difficult at these angles. One plate exposed at 60° was allowed to fade for six days. Upon development the deuteron tracks were found to be almost completely obliterated while the alpha-particle tracks were still visible.

The track lengths in the emulsion were converted to energy by means of the range-energy relation previously reported.[5]

Observations on two unexposed control plates disclosed the presence of a large number of short tracks which were quite evidently alpha-particle tracks. They were isotropically distributed and varied in range from 5 to 20 microns (1.5 to 5.1 Mev). These same tracks were again observed on portions of the reaction plates not exposed to the target and of comparable intensity. The density of tracks in a single group from the target reaction is of the same order of magnitude. The back-

[4] H. Yagoda, *Radioactive Measurements with Nuclear Emulsions* (John Wiley and Sons, Inc., New York, 1949).

[5] Steigert, Toops, and Sampson, Phys. Rev. **83**, 474 (1951).

TABLE I. Energy levels in Mg^{25}. The errors assigned to the energy levels were determined from shifts in the relative values of group energies from plate to plate and are therefore less than the absolute errors in the Q values.

Energy of alpha group (Mev)	Q value (Mev)	Energy levels (Mev) Present work	Previously reported (b)	(c, d, e)
14.37±0.05	6.694±0.01[a]			
13.82±0.05	6.11 ±0.07	0.58±0.02	0.57	
13.56±0.10	5.76 ±0.12	0.93±0.04	0.96	
12.96±0.10	5.07 ±0.12	1.62±0.03	1.63	
12.53±0.20	4.60 ±0.22	2.09±0.05	1.97	
11.98±0.10	3.95 ±0.12	2.74±0.03	2.74	
11.43±0.10	3.33 ±0.12	3.36±0.03	3.36	
10.92±0.10	2.73 ±0.12	3.96±0.04	4.01	
10.81±0.10	2.57 ±0.12	4.12±0.04		
10.14±0.10	1.82 ±0.12	4.87±0.03	4.81	
9.54±0.10	1.13 ±0.12	5.56±0.03	5.48	
9.23±0.10	0.76 ±0.12	5.93±0.03	5.95	
8.33±0.10	−0.29 ±0.12	6.98±0.03		7.40
				7.58
				7.73
7.57±0.15	−1.16 ±0.16	7.85±0.04		7.90
				8.01
				8.16
6.91±0.20	−1.94 ±0.22	8.62±0.05		8.41
				8.59
6.53±0.15	−2.37 ±0.16	9.06±0.04		8.92
				8.96
5.94±0.15	−3.06 ±0.16	9.75±0.04		9.76
5.06±0.15	−4.09 ±0.16	10.78±0.04		
4.09±0.15	−5.20 ±0.16	11.89±0.05		

[a] See reference 7. [b] See reference 8. [c] See reference 8. [d] See reference 9. [e] See reference 10.

ground thus becomes objectionable in measuring track lengths less than 20 microns in length. All track lengths less than 20 microns were therefore rejected from the measurements.

RESULTS AND DISCUSSION

Alpha-particle tracks were measured on ten nuclear emulsion plates which had been exposed to the reaction particles from the aluminum target bombarded by deuterons from the cyclotron. A total of nine thousand tracks were measured. Only those tracks were measured that satisfied conditions of geometry which insured their origin in the center of the target.

The tracks were plotted graphically by means of a triangle method similar to that given by Slätis, Hjalmar, and Carlsson.[6] The peaks were then obtained by numerical integration of the triangles. The half-widths of the triangles used are equal to one unit on the abscissa in Fig. 1. This value was taken as corresponding to the maximum uncertainty in the track length measurement as estimated from the possible distortions, dip, errors due to scale, etc. The half-width used corresponds almost precisely to the half-width of the Po and Th alpha-particle groups used to calibrate the plates.

The data obtained from three plates taken at 90°, 124°, and 60° to the beam are presented in Fig. 1. The total number of tracks measured on each of the plates was 1400, 1200, and 1600, respectively. The peak labeled 0 corresponds to the alpha-particle group leaving Mg^{25} in the ground state.

The energy of each alpha-particle group was obtained from the average value of measurements made on several plates so that even the lowest intensity group corresponds to a total of 300 tracks. The ground-state Q value was taken as 6.694±0.010 Mev from the work of Enge *et al.*,[7] and the energy levels obtained by subtraction from this value. Since the relative differences of the track lengths may be measured much more accurately than the absolute value, the energy levels are correspondingly more accurate. The chief error in the particle energies was from a variation of the deuteron energy from plate to plate which however does not affect the differences in the Q values.

It may be noted that the group labeled VII is a doublet of 180-kev spacing. This appeared in the previous work[2] as a single broad group. Best resolution of this group is obtained on the 124° data.

The energy levels and Q-values obtained by averaging the data for all of the plates are presented in Table I with the previous results shown for comparison. It is seen that the agreement is satisfactory for the two methods of detection.

Nuclear levels calculated from the neutron resonance data of several investigators[8–10] are also given in Table I. These values were obtained by adding the neutron binding energy 7.32±0.02 Mev to the neutron resonance energy (center-of-mass energy). This value was

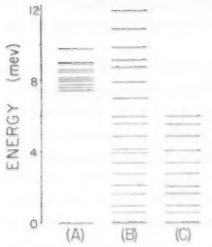


FIG. 2. Comparison of energy levels in Mg^{25}. (A) Neutron levels (see references 8–10), (B) present work, (C) Schelberg, Sampson, and Cochran (see reference 2).

[6] Slätis, Hjalmar, and Carlsson, Phys. Rev. **81**, 641 (1951).

[7] Enge, Buechner, Sperduto, and Van Patter, Phys. Rev. **83**, 31 (1951).

[8] M. R. MacPhail, Phys. Rev. **57**, 669 (1940).

[9] Freier, Fulk, Lampi, and Williams, Phys. Rev. **78**, 508 (1950).

[10] R. E. Fields and M. Walt, Phys. Rev. **83**, 479 (1951).

obtained by using the Q-value 5.094±0.010 Mev for the $Mg^{24}(d,p)Mg^{25}$ reaction,[11] and 2.230±0.007 Mev for the binding energy of the deuteron.[12] If one takes level XII as corresponding to the average of the first six neutron levels, XIII and XIV to the next two doublets, respectively, then the agreement would seem satisfactory. It would appear that at least for the higher excitation of Mg^{25} that the levels obtained in this work may be composed of closely spaced unresolved groups rather than distinct levels.

An energy level diagram comparing the present work with the groups reported by Schelberg[2] and the neutron resonances[8-10] is given in Fig. 2.

[11] Strait, Van Patter, Buechner, and Sperduto, Phys. Rev. **81**, 747 (1951).
[12] R. E. Bell and L. G. Elliot, Phys. Rev. **79**, 282 (1950).

PHYSICAL REVIEW VOLUME 85, NUMBER 2 JANUARY 15, 1952

Angular Distribution in the High Energy Deuteron Photoeffect*

N. AUSTERN†
University of Wisconsin, Madison, Wisconsin
(Received October 5, 1951)

Consideration is given to the possible use of the electric dipole isotropic part of the deuteron photodisintegration cross section as an indicator of weak odd state noncentral forces. This cross section is not very sensitive to weak tensor forces of the usual type, but is strikingly affected by a singular spin-orbit coupling. While for the usual tensor forces the ratio, a/b, of the isotropic to the $\sin^2\vartheta$ term in the cross section varies slowly over a broad energy range, it rises rapidly with energy if the noncentral force is strongly singular. Calculation is done in first Born approximation. A detailed estimate of the magnetic isotropic cross section is included, for this is an unexpectedly large "background" to the electric term which is of interest.

I. INTRODUCTION

THE ground state is the only even parity state of the NP system which is important in deuteron photodisintegration, and its wave function is fairly well known through studies[1] of the many low energy phenomena into which it enters. Considerably less is known about the important states of odd parity. For these, high energy scattering studies[2] have led to a belief in the "even theory," that is, that the exchange dependence of the nuclear forces is such that two nucleons in an odd parity state of relative motion do not interact. The data do not, however, justify the complete exclusion of odd state noncentral interactions, as these are found not to have strong influence on the NP scattering cross section.[2,3] More definite information about them can in principle be obtained from photodisintegration measurements; therefore careful photodisintegration calculations with odd state noncentral forces are desirable. These would supplement the high energy photodisintegration calculations which were recently done for pure central forces.[4,5]

Odd state noncentral interactions do not much affect the total cross section for photodisintegration, but it has long been known from numerical calculation in special cases[6,7] that they do cause one striking qualitative change in the angular distribution. This change consists of the introduction of an isotropic term in the *photoelectric* cross section, which appears if noncentral interactions are effective in the ^3P continuum states. It is produced by interferences between the three ^3P_J states, and is basically a high energy effect, since nucleons in a relative P state cannot interact when the energy is low.

The present work is intended to determine with what sensitivity the isotropic term might actually measure the strength of a weak ^3P state noncentral force. It was begun after accurate low energy measurements of the photodisintegration angular distribution had become available,[8] and when some attempts were being made to perform such experiments at higher energies.[9-13] Energies up to $\hbar\omega = 100$ Mev are considered. This is a range in which it is possible to disregard relativistic and free meson effects[14] and to use the usual phenomenological formalism.[14]

* This work is a portion of the author's doctoral thesis.
† AEC Predoctoral Fellow, now at Cornell University, Ithaca, New York.
[1] For a recent study of this type see Feshbach and Schwinger, Phys. Rev. **84**, 194 (1951). A copy of their work was supplied to the author by Dr. J. Eisenstein.
[2] R. S. Christian and E. W. Hart, Phys. Rev. **77**, 441 (1950).
[3] K. M. Case and A. Pais, Phys. Rev. **80**, 203 (1950).
[4] L. I. Schiff, Phys. Rev. **78**, 733 (1950).
[5] J. F. Marshall and E. Guth, Phys. Rev. **78**, 738 (1950).
[6] W. Rarita and J. Schwinger, Phys. Rev. **59**, 436 and 556 (1941).
[7] T. M. Hu and H. S. W. Massey, Proc. Roy. Soc. (London) **A196**, 135 (1949).
[8] For a summary of recent work see Bishop, Collie, Halban, Hedgran, Siegbahn, du Toit, and Wilson, Phys. Rev. **80**, 211 (1950); Bishop, Halban, Shaw, and Wilson, Phys. Rev. **81**, 219 (1951); Bishop, Beghian, and Halban, Phys. Rev. **83**, 1052(L) (1951).
[9] E. G. Fuller, Phys. Rev. **79**, 303 (1950).
[10] P. V. C. Hough, Phys. Rev. **80**, 1069 (1950).
[11] Phillips, Lawson, and Kruger, Phys. Rev. **80**, 326 (1950).
[12] G. Goldhaber, Phys. Rev. **81**, 930 (1951).
[13] Gibson, Green, and Livesey, Nature **160**, 534 (1949).
[14] See, for example, R. G. Sachs and N. Austern, Phys. Rev. **81**, 705 (1951).

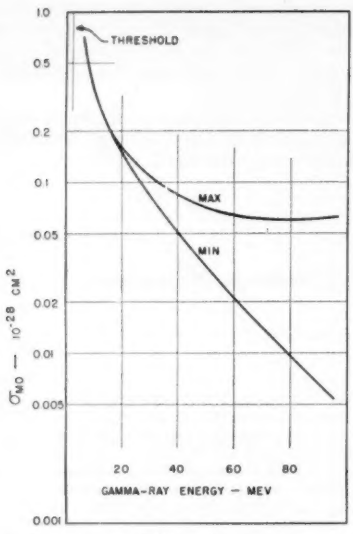


FIG. 1. Photomagnetic disintegration cross section, including only the dipole $^3S \rightarrow {}^1S$ transition, and showing the uncertainty caused by the interaction moment.

In a systematic discussion of possible high energy noncentral forces effects it is important to recognize that for gamma-rays of energy less than 100 Mev the electric dipole radiative interaction always leads to the dominating transitions of the system.[4] These are $^3S_1$ and $^3D_1 \rightarrow {}^3P_{0,1,2}$, and $^3D_1 \rightarrow {}^3F_2$. All other transitions are much weaker and are observable only in so far as they produce interferences with the strongly excited ^3P waves. The $^3F_2$ state can be disregarded entirely; not only is it excited weakly but it also does not contribute any such unique interferences as would be detectable in the presence of other, larger effects.

With pure central forces, electric dipole transitions from the ^3S deuteron ground state lead to ^3P states, and to an angular distribution of the form $\sin^2\theta$. With noncentral forces the angular distribution still cannot possibly[15] have terms of greater complication than $\sin^2\theta$. A term of lesser complication does, however, appear, giving the *electric dipole* angular distribution its most general form, $a_e + b\sin^2\theta$. To the term a_e there must be added the more familiar term, a_m, which comes from photomagnetic transitions to ^1S waves. The entire dipole-induced angular distribution then becomes $a + b\sin^2\theta$, where $a = a_e + a_m$. a_m plays the role of a "background" to the interesting a_e, so its value must be known for the purposes of this paper.

Transitions to continuum triplet waves of even parity can interfere with the ^3P waves and produce fore-aft antisymmetric parts in the angular distribution. Electric quadrupole and magnetic dipole transitions are of this type. With pure central forces only the quadrupole transition $^3S \rightarrow {}^3D$ can occur. The angular distribution becomes $a + b\sin^2\theta(1 + q\cos\theta)$, the disintegration protons favoring the forward hemisphere. The quantity q is given by Schiff,[4] and by Marshall and Guth.[5] The introduction of tensor forces permits the quadrupole transition $^3D \rightarrow {}^3S$, and also magnetic dipole transitions to the coupled continuum $^3S_1 + {}^3D_1$ waves. As it happens, interferences of the altered even waves with the ^3P_J waves do not change q significantly from its value under central forces. These effects will not be discussed further.

In subsequent sections the magnetic cross section a_m is computed first, and then the electric isotropic cross section a_e is computed in three special cases:

(a) No ^3P state nuclear interaction.

(b) A triplet state nuclear interaction of the form

$$V = -(\hbar^2\kappa^2/M)\{\varphi_1[\tfrac{1}{2}(1-x)+\tfrac{1}{2}(1+x)P] + \Gamma S_{12}\varphi_2[\tfrac{1}{2}(1-y)+\tfrac{1}{2}(1+y)P]\}. \quad (1)$$

This interaction is so chosen that for even states it reduces to the expression used by Feshbach and Schwinger in their analysis of the low energy data. P is the Majorana exchange operator. Variation of the parameters x and y gives all possible exchange dependences of V. φ_1 and φ_2 have Yukawa shape, $\varphi = (e^{-\mu r}/\mu r)$. $S_{12} \equiv (3/r^2)(\boldsymbol{\sigma}_1\cdot\mathbf{r})(\boldsymbol{\sigma}_2\cdot\mathbf{r}) - 1$, is the tensor operator. Values of the parameters used in V are given in the appendix.

(c) A very singular $(\mathbf{L}\cdot\mathbf{S})$ interaction. For explicit calculation the expression of Case and Pais[3] is used.

II. PHOTOMAGNETIC CROSS SECTION

The isotropic part of the photomagnetic cross section, a_m, is computed in this section. Effects of the deuteron D function are taken into account, as well as the influence of the spin antisymmetric interaction moment[16] which gives a phenomenological representation of some meson current effects. An $(\mathbf{L}\cdot\mathbf{S})$ interaction of the type of Case and Pais[3] leads to other interaction moments, but these do not have matrix elements with the deuteron ground state.

The photomagnetic matrix element cannot be so easily approximated as the photoelectric matrix element,[17] and this would remain true even were interaction effects not present. Its integrand vanishes less strongly at the origin, by two powers of r, than does that of the electric matrix element. The magnetic matrix element thus tends to be influenced by those parts of the wave function which are most affected by the details of the nuclear potential; and our lack of knowledge of a correct nuclear potential consequently makes the cross section somewhat uncertain. Those other uncertainties which are contributed by the interaction moment are so much greater, however, that the calculations of the present section are only performed with Yukawa well eigenfunctions.[18]

[15] C. N. Yang, Phys. Rev. **74** 764 (1948).

[16] N. Austern and R. G. Sachs, Phys. Rev. **81**, 710 (1951); N. Austern, Phys. Rev. **81**, 307(A) (1951).

[17] H. A. Bethe and C. Longmire, Phys. Rev. **77**, 647 (1950); E. E. Salpeter, Phys. Rev. **82**, 60 (1951).

[18] Descriptions of the various wave functions used in this paper are collected in the appendix.

The magnetic dipole transitions $^3S_1 \rightarrow {}^1S_0$ and $^3D_1 \rightarrow {}^1D_2$ must both be considered. Although the D transition may not be large in itself, it does interfere with the S transition at all energies in such a way as to reduce the apparent value of the magnetic cross section. This can be seen from the differential magnetic dipole photodisintegration cross section:

$$\sigma_m(\theta)=\sigma_{m0}\{1+(5-3\cos^2\theta)J_2^2/4J_0^2 -(3\cos^2\theta-1)\cos(\delta_S-\delta_D)J_2/\sqrt{2}J_0\}, \quad (2)$$

where σ_{m0} is the simple S wave cross section, δ_S and δ_D are the scattering phase shifts, and

$$J_0\equiv\int_0^\infty uu_S\{\tfrac{1}{2}(\mu_N-\mu_P)+\Phi\}dr,$$
$$J_2\equiv\int_0^\infty ww_S\{\tfrac{1}{2}(\mu_N-\mu_P)+\Phi\}dr. \quad (3)$$

The functions u and w are the usual ground-state "radial" wave functions, while u_S and w_S are the corresponding singlet state functions. Φ is the shape function of the spin-antisymmetric interaction moment, which affects both the S and D wave transitions. Equation (2) can be separated into a part with $\sin^2\theta$ angular dependence, and an isotropic part. The former part is then lost experimentally beneath the large $\sin^2\theta$ electric dipole cross section. The latter part, which is the apparent magnetic cross section, is given by

$$a_m=\sigma_{m0}\{1+J_2^2/2J_0^2-\sqrt{2}\cos(\delta_S-\delta_D)J_2/J_0\}. \quad (4)$$

This is smaller than the value σ_{m0}.

In calculating upper and lower bounds on σ_{m0} the spin-antisymmetric interaction moment,

$$\Delta\mathbf{M}=(e\hbar/2Mc)(\boldsymbol{\sigma}_N-\boldsymbol{\sigma}_P)\Phi(r), \quad (5)$$

is assumed as the only correction which must be added to the free nucleon spin magnetic moments, and the form

$$\Phi(r)=\Phi_0(l)\delta(r-l) \quad (6)$$

is once again adopted, $\Phi_0(l)$ being taken from the work of reference 16. σ_{m0} is then found as a function of l, and of photon energy. Figure 1 shows the bounds which are subsequently obtained on σ_{m0} by choosing, at each energy, the largest and smallest values of $\sigma_{m0}(l)$ in the region $l\leqslant 2\times10^{-13}$ cm. The upper limit on the values for l is comparable with the usual ranges for nuclear forces and is probably a little large.

Figure 2 shows a_m as a function of energy. This quantity is derived from σ_{m0} by the prescription of Eq. (4). For the integrals J_0 and J_2 in (4) it is more convenient to employ some definite function Φ rather than the form of Eq. (6). A very reasonable function is $\Phi=\Phi_0 e^{-\mu r}$. For $\mu=8.5\times10^{12}$ cm^{-1}, the value of Φ_0 which gives the correct three-body moment anomaly is $\Phi_0=-0.56$. With this, the ratio a_m/σ_{m0} is computed as a unique function of energy and is used to pass from Fig. 1 to Fig. 2.

A remark is in order about the determination of Φ, the shape function of the interaction moment. Figure 1 shows only that uncertainty which is produced in σ_{m0} because the detailed shape of Φ is unknown. Additional large uncertainties come through lack of exact knowledge of the H^3 wave function and through the extreme idealization involved in the description of the three-body moment anomaly by means of the simple spin-antisymmetric moment. While the H^3 wave function of Pease and Feshbach[19] is certainly much better, the present work is based on the wave function of Avery and Adams.[20] The status of the interaction moment theory does not really seem to justify recalculating all the two-body results on the basis of the better H^3 wave function.

Magnetic quadrupole-magnetic dipole interference might be important at higher energies, but cannot be computed reliably. It would appear through a small fore-aft antisymmetric term in the cross section, which cancels completely when the experimental data are folded about 90° to find a.

III. P WAVE INTERFERENCE—FREE P WAVES

The D wave admixture in the deuteron ground state insures some effect of noncentral forces, even when the ^3P_J waves are entirely free. In particular, the isotropic component, a_e, does not quite vanish for free P waves.

It is convenient to use the differential electric dipole

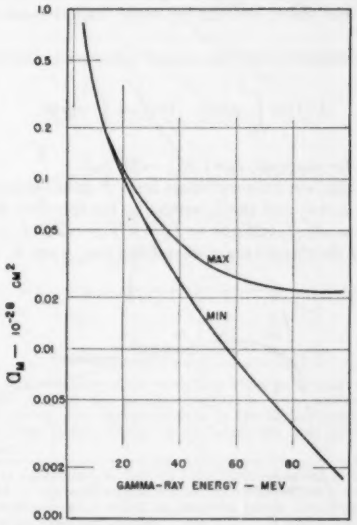


FIG. 2. Isotropic part of the photomagnetic disintegration cross section, including both the dipole transitions $^3S \rightarrow {}^1S$ and $^3D \rightarrow {}^1D$.

[19] R. L. Pease and H. Feshbach, Phys. Rev. **81**, 142(L) (1951); also private communications from Dr. Pease.
[20] R. Avery and E. N. Adams, Phys. Rev. **75**, 1106(L) (1949).

photodisintegration cross section in a slightly different form from that of Rarita and Schwinger:[6]

$$\sigma_e(\theta)d\Omega=\frac{\pi}{3}\frac{e^2}{\hbar c}\frac{M\omega}{\hbar k}\frac{d\Omega}{4\pi}\Big[\Big\{\frac{1}{9}|e^{i\delta_0}I_0-e^{i\delta_2}I_2|^2+\frac{1}{4}|e^{i\delta_1}I_1-e^{i\delta_2}I_2|^2\Big\}+\sin^2\theta\Big\{\frac{1}{4}|e^{i\delta_1}I_1+3e^{i\delta_2}I_2|^2-\frac{1}{15}|e^{i\delta_0}I_0-5e^{i\delta_2}I_2|^2+\frac{I_0^2}{15}\Big\}\Big], \quad (9)$$

where

$$I_0=\int_0^\infty rdrv_0[u-w\sqrt{2}],$$
$$I_1=\int_0^\infty rdrv_1[u+(w/\sqrt{2})], \quad (10)$$
$$I_2=\int_0^\infty rdrv_2[u-(w/5\sqrt{2})].$$

Here $\hbar\omega$ is the gamma-ray energy, $\hbar k$ is the relative momentum of neutron and proton in the continuum state, and the v_J are the ^3P_J state "radial wave functions."

A convenient notation is now introduced. Denote

$$U(f)\equiv\int_0^\infty rdruf, \quad W(f)\equiv\int_0^\infty rdrwf. \quad (11)$$

Then, for example, $I_0=U(v_0)-\sqrt{2}W(v_0)$.

For the "even theory" there is no ^3P wave interaction, so the $\delta_J=0$, and the v_J reduce to the spherical Bessel function, F_1. I_0 reduces to $I_0^0=U(F_1)-\sqrt{2}W(F_1)$, etc., so only the ground-state wave functions, u and w, must be known to compute σ_e. σ_e reduces to

$$\sigma_e(\theta)d\Omega=\frac{\pi}{3(137)}\frac{k^2+\gamma^2}{k}\frac{d\Omega}{4\pi}\Big[\frac{9}{25}W^2(F_1)+\frac{3}{2}\sin^2\theta\times\Big\{U^2(F_1)+\frac{1}{25}W^2(F_1)\Big\}\Big]. \quad (12)$$

$\hbar^2\gamma^2/M$ is the deuteron binding energy, so $\gamma=2.32\times10^{12}$ cm^{-1}.[17] The total cross section is just

$$\sigma_{e0}=(\pi/3(137))(k^2+\gamma^2/k)[U^2(F_1)+(2/5)W^2(F_1)], \quad (13)$$

and

$$(a_e/b)_0=(6W^2(F_1)/25U^2(F_1)+W^2(F_1)). \quad (14)$$

Here the subscript "0" indicates that the results are the ones appropriate to no ^3P wave interaction.

Numerical results have been computed for σ_{e0} and for $(a_e/b)_0$ using the ground-state wave functions which are described in the appendix. $(a_e/b)_0$ is graphed as Fig. 3. Note especially the near-constancy of $(a_e/b)_0$ for $\hbar\omega\gtrsim20$ Mev. This result occurs because u and w have nearly the same shape at small r (as a result of the importance of the S–D cross term in the expectation value of the deuteron potential energy).

Brief consideration is desirable for the effect that noncentral forces in the ground-state configuration might have on the effective range treatment of the low energy photoelectric cross section. There is no room here for a full discussion; that would involve setting up and carrying through an effective range formalism for noncentral forces.[21] The main points of such an analysis can be mentioned: It is only necessary to consider the ground-state S wave, since at low energy the D wave contributes a negligible amount to σ_{e0}. This is illustrated by Fig. 3, $(a_e/b)_0$ being a measure of the D wave contribution. The easily computed zero range photodisintegration cross section is corrected for the nonzero range of forces merely by a renormalization of the ground-state S function. For central forces Bethe and Longmire[17] have shown that the effective range derived from low energy NP scattering experiments can be used very simply to compute the necessary renormalization of the S function. A recheck of the Bethe and Longmire method and of the definition of the effective range in scattering discloses that, while each has to be modified for the presence of the D wave, still, so long as the *experimentally determined* scattering effective range is used to renormalize the S wave, the noncentral forces produce no change in the method. This agrees with a similar remark by Bethe and Longmire.

Thus at very low energy Eq. (13) gives identically the simple central forces result. At high energy the D wave term increases the cross section. For the wave functions used here this increase is roughly constant for energies above about 18 Mev and has the value

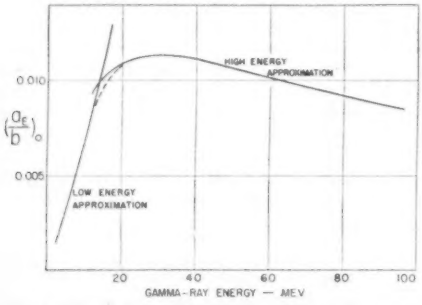


FIG. 3. The ratio of electric isotropic cross section to $\sin^2\theta$ cross section for the case of vanishing ^3P wave forces.

[21] See, for example, J. Schwinger, notes on a course in nuclear physics.

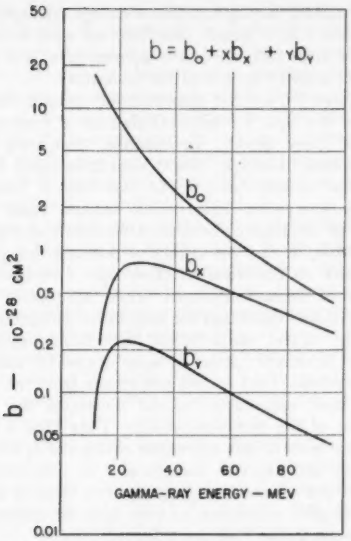


FIG. 4. Coefficient of $\sin^2\theta$ part of electric dipole photodisintegration cross section, showing the first derivatives of the cross section with respect to the exchange parameters for the central and tensor forces. Divide by 4π to get the cross section per steradian.

two percent. If Fig. 3 is again used as a measure of the D wave contribution to $\sigma_{\varepsilon 0}$ the increase of $\sigma_{\varepsilon 0}$ is seen to enter approximately linearly with energy for energies up to 18 Mev.

IV. P WAVE INTERFERENCE—WEAK TENSOR FORCES

In this case Eq. (9) does not reduce to the simple form of (12), for the v_J are affected by the forces. The even theory is considered as the basis for departure, and small deviations from the even theory are computed in first Born approximation. The expression (1) is used for the nuclear potential, where the even theory is obtained if $x=y=0$. Evidently the Born approximation is rigorously correct for very small x and y, and may be expected to give an indication of the effects if x and y become large.

For the odd parity ^3P_J states V becomes

$$V=(\hbar^2\kappa^2/M)\{x\varphi_1+c_J\Gamma y\varphi_2\}, \tag{15}$$

$c_J=-4,\ 2,\ -2/5$, for $J=0,\ 1,\ 2$; and the Schroedinger equation takes the form[22]

$$L_1(v_J)=\kappa^3\{x\varphi_1+c_J\Gamma y\varphi_2\}v_J, \tag{16}$$

[22] Equation (16) is not correct. It is written as if the $^3P_2$ wave were an eigenstate of S_{12}, whereas actually S_{12} couples $^3P_2$ with $^3F_2$. But to the approximation used in this work it is correct to use (16) as given. The complete coupled differential equations for $^3P_2$ and $^3F_2$ are

$$\begin{cases} 0=v_2''-(2/r^2)v_2+[k^2-\kappa^2x\varphi_1+(2/5)\kappa^2\Gamma y\varphi_2]v_2 \\ \qquad -[(6\sqrt{6})/5]\kappa^2\Gamma y\varphi_2 f_2 \\ 0=f_2''-(12/r^2)f_2+[k^2-\kappa^2x\varphi_1+(8/5)\kappa^2\Gamma y\varphi_2]f_2 \\ \qquad -[(6\sqrt{6})/5]\kappa^2\Gamma y\varphi_2 v_2 \end{cases}.$$

where

$$L_1(v_J)=v_J''-(2/r^2)v_J+k^2v_J, \quad L_1(F_1)=0. \tag{17}$$

This differential equation for v_J is solved in terms of the kernel[22]

$$K_1(r,r')=F_1(r_<)G_1(r_>), \tag{18}$$

where $r_<$, $r_>$ are the lesser and greater, respectively, of r, r'; F_1 is the regular solution of $L_1(F_1)=0$, while G_1 is the irregular solution. They are normalized so that the Wronskian $F_1'G_1-F_1G_1'=1$; thus K_1 satisfies

$$L_1(K_1)=-\delta(r-r'). \tag{19}$$

Now to Born approximation

$$\delta_J=-(\kappa^2/k^2)\int_0^\infty F_1^2\{x\varphi_1+c_J\Gamma y\varphi_2\}dr \tag{20}$$

$$\begin{aligned} v_J&=F_1+x\partial v_J/\partial x]_{x=0}+y\partial v_J/\partial y]_{y=0} \\ &=F_1+\delta v_J, \end{aligned} \tag{21}$$

and

$$\begin{aligned} \delta v_J=-\kappa^2\int_0^\infty &K_1(r,r')dr' \\ &\times\{x\varphi_1(r')+c_J\Gamma y\varphi_2(r')\}F_1(r'). \end{aligned} \tag{22}$$

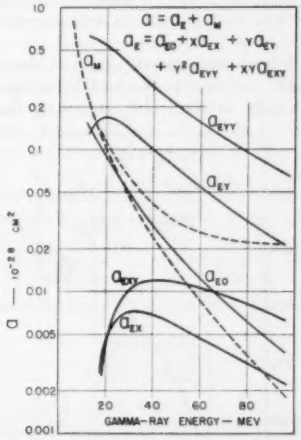


FIG. 5. Coefficients of the quantities which go to make up the isotropic part of the photodisintegration cross section. The dotted curves show the uncertainty in the photomagnetic part of this cross section. Divide by 4π to get the cross section per steradian.

In these equations f_2 is that $^3F_2$ radial function which is strongly coupled to $^3P_2$, and normalized to be $\cos(kr+\delta)$ at $r=\infty$. ϵ is the amplitude with which it properly enters the calculation. $\epsilon\to 0$ as the coupling vanishes.

The first equation above reduces to (16) when the coupling term is not present and reduces rigorously to (16) in the first Born approximation. The coupling term contains as a factor the product of two infinitesimals ϵ and y; thus it vanishes with y to one higher order than the rest of the nuclear potential.

[22] For this formalism see, for example, F. Rohrlich and J. Eisenstein, Phys. Rev. **75**, 705 (1949).

With these expressions the remainder of the calculational procedure is to insert (20) and (21) into (9); perform the necessary integrals; and evaluate numerically. The results of calculation are graphed as Figs. 4, 5, and 6.

Figure 4 shows that the magnitude of the $\sin^2\theta$ part of the cross section, which is nearly the total cross section, depends very strongly on the exchange character of the central potential but hardly at all on the exchange character of the tensor potential. The magnitude of the total cross section is thus a measure of x, the exchange parameter of the central potential. This was indicated before by Marshall and Guth,[5] as well as by other authors.

Figure 5 shows the isotropic cross section, with the photomagnetic background, a_m, included. Study of Fig. 5 shows that a_e is not much influenced by the central potential but does depend strongly on y, the exchange parameter of the tensor force. Above about 15 Mev the tensor forces isotropic component competes strongly with the photomagnetic cross section and dominates it if y is large, i.e., if $y \gtrsim 0.5$. A very favorable energy region for experiments designed to detect the tensor forces isotropic component lies from about 15 Mev to 60 Mev, the lower part of this energy region being the more desirable. At higher energy the photomagnetic cross section becomes unmanageably uncertain.

Figure 6 is an approximate graph of the ratio a_e/b for the case $x=0$ and is obtained by using the even forces value, b_0, in place of b. It is seen that a_e/b is not a rapidly varying function of energy. An experiment to measure a_e/b would, therefore, not need to be performed with monochromatic gamma-rays, and might be well adapted for a betatron laboratory.

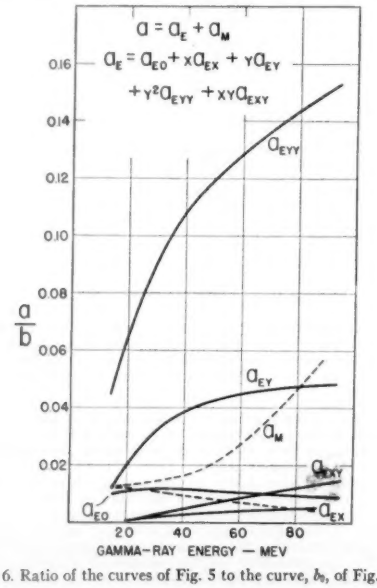


FIG. 6. Ratio of the curves of Fig. 5 to the curve, b_0, of Fig. 4.

Figures 4–6 are not quantitatively reliable for large values of x and y. Exact calculations of some special cases[6,7] have shown, for example, that very much larger tensor isotropic effects sometimes result from a "neutral" theory ($x=y=-1$) than from a "charged" theory ($x=y=1$). This occurs because there is an essential asymmetry between attractive and repulsive potentials, in that an attractive potential has greater influence on the wave function than a repulsive potential of the same strength. While this type of asymmetry is less important for long tailed potentials than for the square wells which have been emphasized before, it should appear to some degree for the cases treated here. That it does not results from the use of first Born approximation and illustrates the inadequacies of the numerical results. Thus Figs. 5 and 6 indicate more or less symmetric changes of a_e for equal positive and negative changes of x or y, so evidently should not be used for large x or y. This is not an unreasonable restriction, as very large deviations from the "even theory" are not expected.

Some care is needed to obtain the maximum information from such an approximate calculation as the one described here, without in the process stepping beyond the approximation. Equation (9) gives $\sigma_e(\theta)$ as a sum of squares. The Born approximation gives the linear correction in x and y to the quantities which are squared, i.e., to the outgoing wave amplitudes. Of course, quadratic and higher corrections also exist. It is clear that, after squaring, the linear terms of σ_e will be known exactly. The quadratic terms of σ_e come partly as squares of the linear corrections to the wave function, which are known, and partly as zero-order second-order cross terms, which are unknown. Where the linear correction to an amplitude is considerably larger than the zero-order amplitude, it is expected that the square of the linear term will be the dominant quadratic addition to σ_e. This situation is nearly met for the noncentral force modification of the isotropic cross section, but not at all for the modifications of the $\sin^2\theta$ cross section. For this reason, the only quadratic terms which are carried are the y^2 and xy terms of a_e.

V. *P* WAVE INTERFERENCE—SINGULAR ($L\cdot S$) COUPLING

For the deuteron the Case and Pais interaction[3] reduces to

$$U=-(2/\hbar)\Lambda(\mathbf{L}\cdot\mathbf{S})\times 6 \text{ Mev}, \tag{23}$$

where

$$\Lambda=(1/\lambda r)[d/d(\lambda r)](e^{-\lambda r}/\lambda r), \quad \lambda=9\times 10^{12} \text{ cm}^{-1}, \tag{24}$$

and for ${}^3P_J$ states,

$$(\mathbf{L}\cdot\mathbf{S})=(\hbar/2)\{J(J+1)-4\}. \tag{25}$$

Here, just as for the tensor interaction, the photodisintegration is computed to Born approximation. This approximation does not now have an automatic domain of validity, because this problem does not involve any parameterization which centers conveniently on vanishing ^3P wave interaction. The justification for using Born approximation is that it is easy, and that it should give some indication of how a very short range ($\mathbf{L}\cdot\mathbf{S}$) interaction might influence the electric dipole photodisintegration.

There is no important correction to the cross section which is linear in the amplitude of U. An important contribution of the ($\mathbf{L}\cdot\mathbf{S}$) interaction is a very large quadratic correction in the isotropic cross section—a correction which rises very rapidly with energy until at 100 Mev it has already surpassed by a factor of ten the largest possible tensor forces contribution. Figure 7 shows a_ϵ and a_ϵ/b_0. While there may be large quadratic corrections to the $\sin^2\theta$ cross section, they are, unfortunately, beyond the scope of the present work.

It should be noted that both the magnitude of the results shown in Fig. 7 and their peculiar energy dependences are due more to the radial shape of the interaction than to its ($\mathbf{L}\cdot\mathbf{S}$) nature. Any strongly singular, noncentral ^3P state interaction must clearly give similar results. The tensor interactions introduced for PP scattering by Christian and Noyes[24] and by Jastrow[25] are other examples. An experiment which determines the angular distribution in photodisintegration by 100-Mev gamma-rays may thus decide whether there were any strongly singular noncentral ^3P state interaction in the NP system. The isotropy of the cross section would be unmistakable.

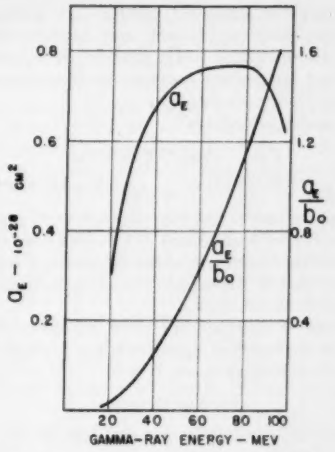


FIG. 7. Isotropic cross section, a_ϵ, if the Case-Pais interaction is present; also its ratio to b_0 of Fig. 4. Divide a_ϵ by 4π to get the cross section per steradian.

VI. CONCLUSIONS

The results of these calculations indicate that a measurement of a_ϵ would be very difficult unless there exist singular noncentral forces in the P states. However, this exception would make the results of such experiments most interesting. In the absence of singular forces, a separation of a_ϵ from the very uncertain a_m becomes nearly impossible unless the odd state strength of the tensor force exceeds 25 percent of its even state strength. The scattering data[2] indicate that the odd state strength is probably not so great.

Aside from the question of magnitude, it is interesting that for nonsingular tensor forces of the usual type there is a broad region of energy in which the magnetic term is not too uncertain and the value of a/b is slowly varying. In this region, a monochromatic source of gamma-rays would not be required. The energy region of interest begins near 20 Mev and may continue as high as 60 Mev, although that upper limit should not be crowded.

The agreement between theory and low energy photodisintegration experiments is well known to be excellent. Measurements of the angular distribution at higher energies have been attempted in several laboratories,[9–13,26] but few have been accurate enough to compare with theory. Goldhaber's result for (a/b) at 7 Mev seems much too large to understand. It is interesting that the experiments have demonstrated the dipole-quadrupole interference effect, and that satisfactory agreement with theory is found. In analyzing any experimental data it should be noted that the very existence of the noncentral forces effects discussed in the present paper vitiates Schiff's[4] remark about the possibility of using the experimental angular distribution for the separation of the electric dipole, electric quadrupole, and magnetic dipole cross sections.

The work was performed with the advice and assistance of Professor R. G. Sachs.

APPENDIX

Wave Functions

All the even state wave functions used in this calculation are eigenfunctions of two nuclear potentials: the triplet potential of Feshbach and Schwinger[1] and a singlet state Yukawa well.

For the triplet potential of Eq. (1) the following parameters are used: $\mu_1=8.45\times10^{12}$ cm^{-1}; $\mu_2=4.71\times10^{12}$ cm^{-1}; $\Gamma=0.224$; $\kappa=11.57\times10^{12}$ cm^{-1}; 2.5 percent D function in the deuteron. That case of Feshbach and Schwinger is chosen which is closest to the set of parameters designated by Pease and Feshbach[19] as giving best results for H^3 binding. The central force range is that derived for low energy PP scattering; the

[24] R. S. Christian and H. P. Noyes, Phys. Rev. **79**, 85 (1950).
[25] R. Jastrow, Phys. Rev. **81**, 165 (1951).
[26] Gibson, Grotdal, Orlin, and Trumpy, Phil. Mag. **42**, 535 (1951).

tensor range is somewhat greater. The phenomena which this potential fits well are: deuteron binding energy; triton binding energy; low energy NP and PP scattering; low energy deuteron photodisintegration; deuteron quadrupole moment.

The singlet potential is[27]

$$V_S=-(\hbar^2K_S^2/M)(e^{-\mu r}/\mu r),$$
$$K_S=10.75\times10^{12}\text{ cm}^{-1},\quad \mu=8.58\times10^{12}\text{ cm}^{-1}.$$

As to the wave functions derived from these potentials: singlet state S and D functions are obtained by numerical integration of the Schroedinger equation, and are used in subsequent numerical integration to compute J_0 and J_2 for σ_m.

For the ^3S ground-state wave the best Hulthén function, an excellent approximation, is fitted to the Feshbach-Schwinger curve. It is

$$u=N(e^{-\gamma r}-e^{-\xi r}),$$

where $\gamma=2.316\times10^{12}$ cm^{-1}, $\xi=13.36\times10^{12}$ cm^{-1}, and $N=(7.76\times10^{12}\text{ cm}^{-1})^{\frac{1}{2}}$.

For the ^3D ground state wave several different procedures are used. First, for magnetic calculations, the Feshbach-Schwinger curve is used directly in numerical integration. Two approximate analytic expressions are also used, which are fitted to the numerical curve. A low energy approximation is

$$w\approx M'(1-e^{-\alpha r})e^{-\gamma r}[1+(3/\gamma r)+(3/\gamma^2r^2)],$$

which has the exact asymptotic form. $\alpha=3.60\times10^{12}$ cm^{-1}.

$$M'=(0.0100\times10^{12}\text{ cm}^{-1})^{\frac{1}{2}}.$$

A high energy approximation is

$$w\approx M\{(1-Ar)e^{-\xi r}-Be^{-\eta r}\},$$

where $\xi=3.06\times10^{12}$ cm^{-1}, $A=0.1385\times10^{12}$ cm^{-1}, $B=1.140$, $\eta=11.04\times10^{12}$ cm^{-1}, and $M=(0.455\times10^{12}\text{ cm}^{-1})^{\frac{1}{2}}$.

[27] J. M. Blatt and J. D. Jackson, Phys. Rev. **76**, 18 (1949).

Phase Shifts

The phase shifts δ_S and δ_D of Sec. II were determined first by approximate formulas, and the approximations then improved in the subsequent numerical integration of the wave equation. It may be of some interest to present the Born approximation formula for δ_D, if the potential is V_S, above. It is

$$\delta_D=\frac{K_S^2}{4k\mu}\left\{\left[1+\frac{3}{2}\left(\frac{\mu}{k}\right)^2+\frac{3}{8}\left(\frac{\mu}{k}\right)^4\right]\log\left(\frac{\mu^2+4k^2}{\mu^2}\right)-3\left[1+\frac{1}{2}\left(\frac{\mu}{k}\right)^2\right]\right\}.$$

PHYSICAL REVIEW VOLUME 85, NUMBER 2 JANUARY 15, 1952

The Absorption of Gamma-Rays from Co^{60}

SAKAE SHIMIZU, TETSUYA HANAI, AND SUNAO OKAMOTO
Nuclear Research Laboratory, Department of Physics, University of Kyoto, Kyoto, Japan
(Received September 7, 1951)

Measurements of the absorption of gamma-rays from Co^{60} (1.17 and 1.33 Mev) have been made in 27 elements. In order to exclude the errors due to secondary radiations which might be produced in neighboring objects and in the absorbers, particular precautions were taken with respect to the geometrical arrangement of apparatus. The absorption coefficients measured for the elements whose atomic numbers are less than $_{52}$Te show reasonable agreement with those calculated according to existing theories. However, it is noted that the results with $_{73}$Ta, $_{74}$W, $_{78}$Pt, $_{79}$Au, $_{80}$Hg, and $_{81}$Tl are 3 percent to 5.5 percent less than the theoretical values. It seems improbable that the disagreements observed in these elements may be assigned to experimental causes. If the entire deviation were assigned to inaccuracy in theoretical knowledge, it would be reasonable to attribute it to some insufficiency in the Klein-Nishina theory of the Compton effect for this energy of gamma-rays. But further investigation should be undertaken to ascertain the fact.

I. INTRODUCTION

SINCE the application of Co^{60} has rapidly increased in various fields of science, it becomes important to know with greater accuracy the absorption coefficients of the gamma-rays from this radioisotope (1.17 and 1.33 Mev) in various elements. The absorption of gamma-rays in matter may be attributed to the combination of four separate effects, namely the photoelectric effect, the Compton effect, pair production, and the photonuclear reaction. Photonuclear reactions seem to be generally improbable in the energy range below several Mev except for a few nuclei. The absorption due to the photoelectric effect has been theoretically estimated by many workers,[1] and that due to the Compton effect has been formulated by Klein and Nishina,[2] while pair production has been theoretically discussed by Dirac and others.[3] A summary of most of these theories, which give the knowledge of absorp-

[1] F. Sauter, Ann. Physik **9**, 217 (1931); **11**, 454 (1931); H. Hall, Phys. Rev. **45**, 620 (1934); H. Hall and W. Rarita, Phys. Rev. **46**, 143 (1934); J. G. Jaeger and H. R. Hulme, Proc. Roy. Soc. (London) **148**, 708 (1935); Hulme, McDougall, Buckingham, and Fowler, Proc. Roy. Soc. (London) **149**, 131 (1935).
[2] O. Klein and Y. Nishina, Z. Physik **52**, 853 (1928).
[3] P. Dirac, Proc. Cambridge Phil. Soc. **30**, 150 (1934); W. Heisenberg, Z. Physik **90**, 209 (1934); H. Bethe and W. Heitler, Proc. Roy. Soc. (London) **146**, 83 (1934); W. Furry and J. R. Oppenheimer, Phys. Rev. **45**, 245 (1934).

tion of gamma-rays in matter, has been given by Heitler.[4]

Several measurements of absorption of gamma-rays from some radioisotopes and comparisons with the theoretical values have already been published. Among these reports it is noted that Cork and Pidd[5] found apparent discrepancies between observed and computed values of absorption coefficients in lead and copper using gamma-rays from Zn^{65} (1.14 Mev), Na^{24} (2.76 Mev), and Co^{60}. However, the result of the absorption measurement of Na^{24} gamma-rays in lead obtained by Groetzinger and Smith[6] substantiated the theoretical value. Furthermore, Davisson and Evans[7] have recently measured accurate absorption coefficients of gamma-rays from I^{131} (0.367 Mev and others), Cu^{64} (annihilation radiation), Mn^{54} (0.835 Mev), Co^{60}, Zn^{65}, and Na^{24}, and found that the values obtained by their experiments showed agreement within 0.5 percent to 2 percent with theory. However, an anomalous absorption coefficient 5 percent less than the theoretical value with tantalum absorber and Zn^{65} and Co^{60} sources found by their measurements seems to need further investigation.

In the present work, undertaken to obtain more information on the absorption of gamma-rays in many elements, we made absorption measurements for 27 elements using the gamma-rays from Co^{60} with a special arrangement of the apparatus; and we compared the results obtained with theoretical values.

II. APPARATUS AND EXPERIMENTAL PROCEDURE

The source of Co^{60} used in the present experiment was prepared from radioactive cobaltous chloride solution, which was supplied from Oak Ridge last year by courtesy of the United States Atomic Energy Commission. Evaporation residuum of this solution of about 2 millicuries equivalent intensity, sealed in a thin glass sphere of 3 mm in diameter, was used as a gamma-ray source hung from the ceiling of the room in which the experiment was performed (Fig. 1).

As a gamma-ray detector we used an end-window G-M counter, 2 cm in diameter with an effective length of about 3 cm, filled with a mixture of argon and alcohol vapor. In practice, however, we used this counter with a lead plate of 1.2 mm thickness placed just upon a mica window for the reason described below. The counter was connected to a high speed scale-of-100 recording circuit.

In the geometrical arrangement of apparatus involved in the measurement of gamma-ray absorption, it is essential to minimize secondary photons and electrons from neighboring bodies and from the absorber itself as far as possible, since such scattered radiations may reach the detector. In work so far published, various precautions were taken in this respect; namely, the adequate collimation of gamma-rays was generally made by placing the source in a deep hole situated in a solid lead block, the detector was surrounded on all sides by lead except for the side facing the source, and the solid angle subtended by the detector at the source and absorber was minimized. However, it should be noted that the lead block holding the source and the collimation lead may produce scattered radiations.

Taking into account these conditions, we arranged the apparatus as shown schematically in Fig. 1. A gamma-ray source without any lead block was hung from the ceiling of the room at 80 cm above the detector, an end-window G-M counter; to define the path of the gamma-rays falling upon the detector, a lead cylinder, *B*, 2.7 cm in inner diameter and 9.8 cm in outer diameter, with a height of 11.2 cm, was placed just above the counter. In order to prevent the secondary electrons. which might be produced at the absorber and the inner wall of the lead cylinder, from reaching the counter, and to count only the gamma-ray quanta

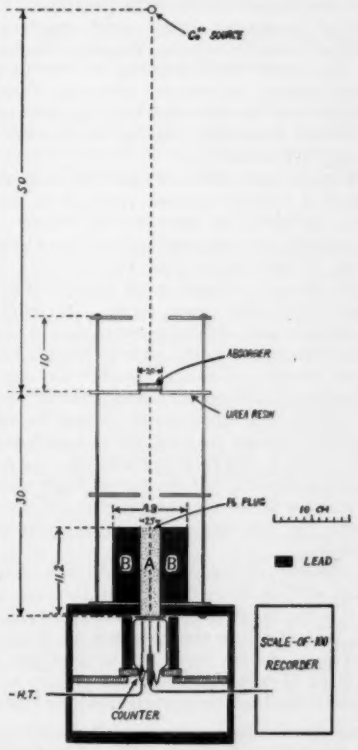


FIG. 1. Geometrical arrangement of apparatus to measure absorption coefficients.

[4] W. Heitler, *The Quantum Theory of Radiation* (Oxford University Press, London, 1936).
[5] J. M. Cork and R. W. Pidd, Phys. Rev. **66**, 227 (1944); J. M. Cork, Phys. Rev. **67**, 53 (1945).
[6] G. Groetzinger and L. Smith, Phys. Rev. **67**, 53 (1945).
[7] C. M. Davisson and R. D. Evans, Phys. Rev. **81**, 404 (1951).

TABLE I. Absorption coefficient of gamma-rays of Co^{60} in lead, as measured with different distances between source, absorber, and counter.

Distance between source and counter (cm)	Distance between absorber and source (cm)	Absorption coefficient μ(cm^{-1})
80	30	0.6426±0.0028
80	40	0.6376±0.0031
80	50	0.6382±0.0047
150	30	0.6382±0.0058
150	40	0.6410±0.0047
150	50	0.6375±0.0044

passing through the canal of the cylinder, the mica window of the counter was covered with a thin lead plate of 1.2 mm thickness and the side wall of the counter was sufficiently surrounded by lead as shown in Fig. 1.

The measurement procedure with this geometry was as follows: for each absorber first we measured the count, c_1, with a lead plug, A, which filled exactly the canal of the lead cylinder, B, and then the count, c_2, without this plug. The difference of counts with and without the plug, c_2-c_1, should be proportional to the number of gamma-ray quanta which penetrate the absorber and pass through the canal, since the count c_1 seems practically to be due only to the scattered radiation reaching the counter sidewards. Therefore, we took c_2-c_1 as the true count of the gamma-rays for each absorber. In practice, when no absorber was used, c_1 was only 1.2 percent of c_2.

With this geometry and procedure[8] we could exclude the effect of scattered photons produced in the surrounding matters in the room, and of secondary electrons produced in the absorber and in the lead block, B. However, it was feared that the Compton singly scattered photons produced in the absorber itself and the lead block might cause counts. To clarify this point we made some absorption measurements with the procedure above mentioned and with different relative distances between the source, absorber, and detector, which resulted in different scattering angles with respect to the detector. The experimental values of the absorption in lead obtained by these test measurements are shown in Table I. The results agreed within the experimental error, showing that the effect of singly scattered photons falling on the counter was negligible with this geometry. For this reason we adopted the geometry shown in Fig. 1.

The absorbers of most elements, 3 cm in diameter, were placed 30 cm above the detector. However, in the cases of a few elements such as Mn, Ta, W, Pt, Au, and Tl, for which 3 cm absorbers were not obtainable, we used plates of about 2.5×2.5 cm as absorbers. These were placed 10 cm above the normal position so that their projections from the source covered the window of the counter. This alternation of position of absorbers would not affect the final results as shown by test measurements above mentioned. The absorption in metallic mercury was measured by filling a glass cylinder whose absorption when empty was known. For each element, several absorbers of different thicknesses were used, so that we might be able to obtain several points on the absorption curve until the count of gamma-rays decreased to about one-third of that without absorber. Generally, we paid no attention to the presence of impurities in absorber elements, since the quantitative chemical analysis of the absorbers used showed that the purities of most elements were at least more than about 99 percent. Densities of most absorbers were precisely determined from their weights measured in both air and pure water, while those of a few samples such as carbon and calcium were determined from measurements of volume and weight in air.

The absorption coefficients were calculated from the observed transmission and measured absorber densities, using the method of least squares.

[8] Similar geometry and procedure were used by Dr. Y. Uemura for the absorption measurements of high energy gamma-rays from Li(p,γ) and F(p,γ) reactions [Bull. Inst. Chem. Research, Univ. of Kyoto **22**, 18 (1950)].

III. RESULTS AND DISCUSSION

Experimental results obtained are summarized in Table II. In the fifth column of the table are listed atomic absorption coefficients expected theoretically.

In the computation of theoretical values, for the energies of the gamma-rays from Co^{60} we used the precise values measured by Lind, Brown, and DuMond[9]

TABLE II. Measured absorption coefficients of gamma-rays from Co^{60} in 27 elements compared with those computed theoretically.

Element	Linear abs. coef. μ (cm^{-1})	Mass abs. coef. μ_m (cm^2/g)	Atomic abs. coef. Experimental σ_{exp} (barn)	Atomic abs. coef. Theoretical σ_{th} (barn)
$_6$C	0.0884±0.0005	0.0589±0.0003	1.175±0.006	1.133
$_{12}$Mg	0.0994±0.0008	0.0572±0.0005	2.307±0.020	2.267
$_{13}$Al	0.1433±0.0011	0.0529±0.0004	2.378±0.018	2.457
$_{16}$S	0.1121±0.0004	0.0567±0.0002	3.021±0.011	3.026
$_{20}$Ca	0.0895±0.0011	0.0584±0.0007	3.887±0.046	3.788
$_{22}$Ti	0.2286±0.0025	0.0527±0.0006	4.189±0.046	4.171
$_{25}$Mn	0.3792±0.0096	0.0513±0.0013	4.675±0.119	4.750
$_{26}$Fe	0.4076±0.0037	0.0519±0.0005	4.826±0.044	4.943
$_{27}$Co	0.4469±0.0020	0.0512±0.0002	5.013±0.022	5.137
$_{28}$Ni	0.4881±0.0022	0.0546±0.0002	5.324±0.024	5.333
$_{29}$Cu	0.4638±0.0064	0.0516±0.0007	5.506±0.076	5.529
$_{30}$Zn	0.3662±0.0027	0.0513±0.0004	5.565±0.041	5.726
$_{34}$Se	0.2154±0.0017	0.0501±0.0004	6.572±0.053	6.523
$_{42}$Mo	0.5268±0.0038	0.0514±0.0004	8.192±0.059	8.179
$_{47}$Ag	0.5305±0.0032	0.0505±0.0003	9.046±0.055	9.283
$_{48}$Cd	0.4382±0.0049	0.0505±0.0006	9.418±0.104	9.512
$_{50}$Sn	0.3621±0.0035	0.0496±0.0005	9.779±0.096	9.971
$_{51}$Sb	0.3349±0.0026	0.0500±0.0004	10.11 ±0.08	10.21
$_{52}$Te	0.3042±0.0023	0.0487±0.0004	10.33 ±0.08	10.46
$_{73}$Ta	0.8990±0.0202	0.0535±0.0012	16.06 ±0.36	16.56
$_{74}$W	1.011 ±0.005	0.0527±0.0003	16.10 ±0.08	16.90
$_{78}$Pt	1.167 ±0.023	0.0543±0.0011	17.61 ±0.34	18.39
$_{79}$Au	1.074 ±0.005	0.0555±0.0003	18.17 ±0.09	18.78
$_{80}$Hg	0.7493±0.0095	0.0553±0.0007	18.41 ±0.23	19.18
$_{81}$Tl	0.6322±0.0017	0.0542±0.0001	18.53 ±0.05	19.60
$_{82}$Pb	0.6426±0.0028	0.0566±0.0002	19.46 ±0.09	19.97
$_{83}$Bi	0.5658±0.0025	0.0579±0.0003	20.10 ±0.09	20.42

[9] Lind, Brown, and DuMond, Phys. Rev. **76**, 1838 (1949).

with the use of the two-meter focusing curved-crystal gamma-ray spectrometer, namely 1.3316±0.0010 Mev and 1.1715±0.0010 Mev with equal intensities.

For the photoelectric absorption, we used the theory developed by Hulme, McDougall, Buckingham, and Fowler,[10] who gave their results in the form of a graph showing a calculated function of the photoelectric absorption coefficient per atom over a range of atomic number and quantum energies. Since the procedure of their theoretical calculation was very complicated, values of the photoelectric absorption coefficient per atom were deduced by graphical interpolation from their results for the elements and gamma-ray energies used by us.

For the computation of the atomic cross section due to Compton scattering we used the well-known Klein-Nishina formula:

$$\sigma_{\text{Comp}}=2\pi r_0^2 Z\left[\left(\frac{1+\alpha}{\alpha^2}\right)\left(\frac{2(1+\alpha)}{1+2\alpha}-\frac{1}{\alpha}\ln(1+2\alpha)\right)+\frac{1}{2\alpha}\ln(1+2\alpha)-\frac{1+3\alpha}{(1+2\alpha)^2}\right], \quad (1)$$

where $r_0=e^2/m_0c^2$, the classical electron radius, Z is the atomic number, and α is the gamma-ray energy in unit of m_0c^2.

The cross section per atom for the pair production for gamma-ray energies less than $\alpha=10m_0c^2$ was expressed by Hirschfelder and Adams[11] in the following relation,

$$\sigma_{\text{pair}}=0.2545r_0^2(Z^2/137)(\alpha-2.332). \quad (2)$$

We applied this relation to the computation of the cross section for the pair production, because this expression gave good agreement with the numerical values cited by Heitler.[12] However, since this relation was not valid for the gamma-ray energies less than $\alpha=2.332m_0c^2$, we took the cross section for the lower energy component of gamma-rays from Co^{60} ($\alpha=2.2933m_0c^2$) to be negligibly small for any element used, namely to be $\sigma_{\text{pair, 1.17}}=0.0000\times10^{-24}$ cm^2.

The total cross section for each component of the gamma-rays used is given by the sum,

$$\sigma_{\text{tot, 1.17}}=\sigma_{\text{phot, 1.17}}+\sigma_{\text{Comp, 1.17}}+\sigma_{\text{pair, 1.17}}$$

for 1.1715-Mev photon,

and

$$\sigma_{\text{tot, 1.33}}=\sigma_{\text{phot, 1.33}}+\sigma_{\text{Comp, 1.33}}+\sigma_{\text{pair, 1.33}}$$

for 1.3316-Mev photon.

Further, taking into account the facts that these components have equal intensities and that the gamma-ray counter used has the different counting efficiency for each of these photons, the theoretical total cross section per atom for the Co^{60} gamma-radiation, to be compared with the experimental value obtained by the present experiment, should be calculated by the following relation,

$$\sigma_{\text{th}}=(R\sigma_{\text{tot, 1.17}}+\sigma_{\text{tot, 1.33}})/(R+1), \quad (3)$$

where $R=\epsilon_{1.17}/\epsilon_{1.33}$ is the ratio of counting efficiencies of the counter for each component of the gamma-rays used. In the present computation $\epsilon_{1.17}$ and $\epsilon_{1.33}$ were taken to be 1.00 percent and 1.11 percent respectively, according to the experimental values determined by the work of Bradt *et al.*[13] for a lead counter of 1.0 mm thickness.

The atomic absorption coefficients computed theoretically by these procedures are listed in the fifth column of Table II, and measured and theoretical values plotted against atomic number are shown graphically in Fig. 2. The solid curve in this figure shows the result expected theoretically.

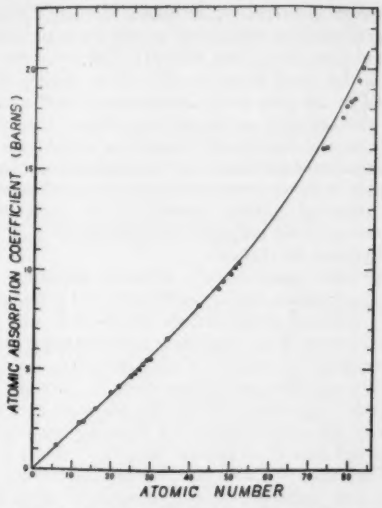


FIG. 2. Experimental and theoretical values of the absorption coefficient per atom plotted against atomic number Z. The solid curve indicates the result expected theoretically.

The results obtained with the elements whose atomic numbers are less than $_{52}$Te show agreement within 0.2 percent to 2.5 percent with theory; and with some elements, our results are in agreement within the experimental errors with those of other workers who have used the same source.[5,7,14] However, it is noted that the results with $_{73}$Ta, $_{74}$W, $_{78}$Pt, $_{79}$Au, $_{80}$Hg, and $_{81}$Tl are 3 percent to 5.5 percent less than the values

[10] Hulme, McDougall, Buckingham, and Fowler, Proc. Roy. Soc. (London) **149**, 131 (1935).
[11] J. O. Hirschfelder and E. N. Adams, II, Phys. Rev. **73**, 863 (1948).
[12] See reference 4, 2nd edition, 1944, Appendix II, p. 259.
[13] Bradt, Gugelot, Huber, Medicus, Preiswerk, and Scherrer, Helv. Phys. Acta **19**, 77 (1946).
[14] W. V. Mayneord and A. J. Cipriani, Can. J. Research **A25**, 303 (1947).

computed theoretically, as shown in Fig. 2. These disagreements are apparently outside the experimental error of about 2.5 percent. We have checked the possible reasons for these deviations. The error which may be caused by the presence of impurities in absorbers used was checked, but we found immediately that their effect on the final results should be masked by the statistical error involved in the counting measurements, since the impurities were generally less than 0.5 percent. The chemical analysis showed, for instance, 99.98 percent purity for tungsten, 99.99 percent for gold, and 99.91 percent for thallium.

The other possible source of errors responsible for the disagreement above mentioned is the Compton singly scattered radiation from the absorber itself and the inner wall of the canalizing lead slit placed before the counter. An analysis of such secondary radiation reaching the detector has been made by Tarrant[15] and Davisson and Evans.[7] Under the geometry here considered, the correction for this secondary radiation is small and may be evaluated using the formula given by Davisson and Evans for the difference between the true and apparent atomic absorption coefficient, namely,

$$\sigma_{\text{true}}-\sigma_{\text{app}}=\pi r_0^2 Z\theta_0^2[1-(\theta_0^2/12)(9\alpha+4)], \quad (4)$$

where θ_0 is the maximum angle of scattering for radiation entering the center of the detector.

From Eq. (4) we can estimate the effect of the scattered radiation from the absorber on the final result. For the Co^{60} radiation α is about 2.4 and the maximum angle of scattering in the present geometry is at most 0.0803 radian, so that the maximum scattering correction is about $0.0016\times10^{-24}Z$ cm^2. The amount of such a small correction for each of the absorber elements is nearly equal to or less than the statistical error due to the counting measurements. Therefore, the effect of the secondary radiation from the absorber may be neglected.

To estimate further the effect of the secondary radiation from the inner wall of the lead canal, we made transmission measurements with the source and absorber in different positions, thus varying the solid angle subtended by the inner wall at the source. The measured values show agreement within the statistical errors as shown in Table I, telling us that such secondary radiation is too minute to have any appreciable effect on the final results. From these considerations it can be affirmed that the deviations from theory observed with some elements should not be assigned to experimental causes, particularly not to the effect of the Compton scattered radiation.

For the gamma-rays from Co^{60} the absorption in the elements with which deviations from theory were found is primarily due to the Compton effect. If the entire deviation observed were assigned to the theoretical insufficiency of the photoelectric absorption, then a reduction of the absorption coefficient by about one-fourth would be required. This seems unlikely since reasonable agreement between the experimental and theoretical values was found in the other elements we examined. And since the absorption due to the pair production is vanishingly small at these photon energies, it is impossible to explain the observed discrepancies even by reducing this absorption to zero. Therefore, if the deviations found in $_{73}$Ta, $_{74}$W, $_{78}$Pt, $_{79}$Au, $_{80}$Hg, and $_{81}$Tl are real, it may be more reasonable to attribute them to some insufficiency in the Klein-Nishina theory of the Compton effect for photons of about 1.2 Mev. However, in order to ascertain this fact, further tests should be made with a stronger source of Co^{60} or Zn^{65}.

In conclusion, the authors wish to express their sincere appreciation for valuable discussions given by Professor K. Kimura and Dr. Y. Uemura. We have to thank sincerely the United States Atomic Energy Commission and Dr. H. C. Kelly, former chief of the Scientific and Technical Division, ESS–GHQ–SCAP, and the Scientific and Technical Administration Committee of Japan for introducing radioisotopes into Japan from Oak Ridge, Tennessee, U.S.A. Thanks are also due to Professor T. Okada and Professor K. Nishihara of the University of Kyoto, and to Kobe Kogyo Company, Ltd., and Mitsubishi Electric Manufacturing Company, Ltd. for the loan and the kind presents of some samples of pure elements, which enabled us to perform the work. This work was financially supported by the special fund of the Ministry of Education for the Advancement of Science.

[15] G. T. P. Tarrant, Proc. Cambridge Phil. Soc. **28**, 475 (1932).

The Energy Spectrum of Primary Cosmic Radiation*

M. F. KAPLON, B. PETERS, H. L. REYNOLDS,† AND D. M. RITSON
University of Rochester, Rochester, New York

(Received August 27, 1951)

Techniques have been developed by which the energy spectrum of heavy primary nuclei can be measured in the energy range of 10^{9}–10^{11} ev/nucleon and of both protons and α-particles in the energy range from 10^{11} to 10^{13} ev/nucleon. It is shown that even the heaviest nuclei in the primary radiation ($Z\sim26$) have lost their orbital electrons before entering the atmosphere, which supports the assumption that they do nor proceed from the solar corona directly to the earth. It is also shown that the velocity spectrum of all primaries of charge $Z>1$ is independent of the mass of the nuclei, and that it differs only slightly from that of the proton component in the energy interval between 3×10^{9} and 10^{10} ev/nucleon.

INTRODUCTION

PRIMARY cosmic radiation is known to consist of a mixture of nuclei whose atomic numbers range from $Z=1$ to $Z\sim26-30$.[1-4] The relative abundance of different elements in the cosmic-ray beam seems to parallel closely the relative abundance of chemical elements in the visible universe.

In order to gain further insight into the possible mechanism by which these particles may have been accelerated, it seems useful to study separately the energy spectrum of different components of the primary beam.

Data on the energy spectrum of the total primary radiation have been obtained by observing total counting rates at the top of the atmosphere.[5] Since at any given latitude over 80 percent of the primary particles are protons, these investigations yield essentially the energy spectrum of the proton component in the range between 1 and 15 Bev. Some measurements on the energy spectrum of heavier primaries have been obtained previously by measuring the flux of various components of the primary radiation at different latitudes. In Chapter I of this paper we shall present some results on the energy spectrum obtained from nuclear interactions of primary particles. The energy region covered by these measurements includes most of the latitude-sensitive part of the spectrum (<6.8 Bev/nucleon) and extends beyond it up to energies of about 50 Bev/nucleon.

These measurements together with new measurements in the nonrelativistic region of the spectrum (Chapter II) and additional flux data (Chapter III) seem to indicate that the velocity spectrum of all components of the primary beam from helium to iron is identical and can be represented by an integral spectrum of the form $N(\epsilon)=(K/(1+\epsilon)^{1.35\pm0.15})$, where ϵ is the kinetic energy per nucleon in Bev and K is a number which for each charge component is independent of energy.

Chapter IV contains a comparison of data obtained from the latitude effect with those obtained from the range and from nuclear interactions of the primary particles. This comparison provides conclusive evidence that even primary particles as heavy as iron arrive from outer space completely or very nearly completely stripped of their electronic shells. They must have lost their electrons either in the process of acceleration or in their subsequent traversal of interstellar matter. An estimate for the ionization of hydrogen-like atoms by electrons would indicate a minimum age for iron nuclei of the order of (300 years)/ρ, where ρ is the number of of protons and electrons per cc along the trajectory. This corresponds to a minimum amount of material traversed of 12 μg or about 200 times more than would be traversed by a particle proceeding directly from the solar corona to the earth.

In the course of these investigations we have obtained values for the mean-free path for nuclear collisions in brass (Chapter V) both for nuclei of charge $Z\geqslant6$ and energy above 1-Bev/nucleon and for α-particles of energy above 10 Bev. When combined with values previously obtained for the mean-free path in air and glass, it appears that the cross section for nuclei of atomic weight between helium and iron, in materials ranging in atomic weight from air to brass, can be adequately represented by the formula given previously for a more restricted range of atomic weights.[2,3]

In Chapter VI we shall briefly describe some measurements now in progress which permit determination of the flux of protons and α-particles in the energy range between 10^{11} and 10^{13} ev per nucleon. Incomplete results indicate that in this energy region the ratio of primary α-particles to protons is similar to the ratio in the latitude sensitive part of the beam.

CHAPTER I. MEASUREMENTS OF THE ENERGY OF HEAVY PRIMARIES BASED ON NUCLEAR COLLISIONS

(1) Description of the Apparatus

All the results reported in this paper are based on two high altitude flights carried out by the General

* This research was supported by the AEC.

† AEC Predoctoral Fellow, now at Oak Ridge National Laboratory, Oak Ridge, Tennessee.

[1] P. Freier *et al.*, Phys. Rev. **74**, 213 (1948).
[2] H. Bradt and B. Peters, Phys. Rev. **77**, 54 (1950).
[3] H. Bradt and B. Peters, Phys. Rev. **80**, 943 (1950).
[4] B. Peters, *Progress in Cosmic Ray Physics* (North Holland Publishing Company, Amsterdam, 1951), Vol. 1, Chap. IV.
[5] J. R. Winckler *et al.*, Phys. Rev. **79**, 656 (1950).

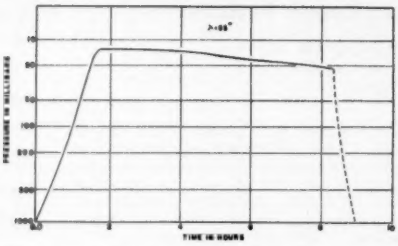


FIG. 1. Time-altitude curve for Minnesota balloon flight.

Mills Aeronautical Laboratory. One flight was made at Minnesota (geomag. latitude $\lambda = 55°$) on September 24, 1950, the other at White Sands, N. M. ($\lambda = 41.7°$) on February 12, 1951. Figures 1 and 2 show the time altitude curves for these flights.

In both experiments the same apparatus was used. It consisted of the box shown in Fig. 3. This box is made of 20 brass plates 3 mm thick and accurately spaced by a brass frame. Below each brass plate a nuclear G-5 emulsion of 100μ thickness mounted on a glass plate 4 in.$\times$6 in. and 1.3 mm thick was inserted. Great care was taken to insure that the plates were accurately aligned with respect to each other and evenly spaced. Six 250μ emulsions (G-5, NTB3, C2, NTA) were placed on top of the assembly. The stack was flown with the plates in a horizontal position.

(2) Collisions of Heavy Nuclei

Stars produced by energetic heavy nuclei differ in some respects from those produced by protons or α-particles. If a proton or α-particle hits a target nucleus the resulting fragmentation of the target nucleus gives rise to a variety of tracks which may roughly be classified as follows:

(a) Shower particles causing "light tracks." These consist of singly charged particles of high energy with near minimum ionization. Fowler[6] has shown that more than 75 percent of these particles are π-mesons, the remainder mostly protons. They appear only in collisions involving energies in excess of $\sim$1 Bev and increase in multiplicity as the energy increases. The tracks are usually collimated in the forward direction, the degree of collimation increasing with multiplicity.

(b) Fragments producing "grey tracks." These consist mostly of protons in the energy range from about 20 Mev to 200 Mev with a small admixture of heavier nuclei. The tracks are not collimated but are preferentially emitted in the forward hemisphere. These particles are probably the result of the nucleonic cascade started by the incident proton.

(c) Evaporation tracks consisting of about equal numbers of protons and α-particles with energies of a few Mev and some heavier fragments whose charge rarely exceeds that of a carbon nucleus. The angular distribution of these tracks is spherically symmetric.

(d) Recoil fragments. This is the part of the target nucleus which has escaped destruction in the collision and has cooled down by the evaporation of α-particles, protons and neutrons. A recoil fragment is almost always present if the target nucleus was heavy. It produces very short tracks rarely exceeding a few microns in length.

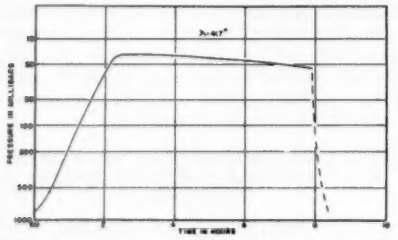


FIG. 2. Time-altitude curve for White Sands balloon flight.

[6] P. H. Fowler, Phil. Mag. 41, 169 (1950).

If the incident nucleus is heavy, an identical set of fragments will be produced in the rest system of the projectile. The recoil fragment will in the laboratory system appear as a heavy nucleus proceeding in the direction of the incident particle with undiminished velocity. The evaporation tracks will appear as a strongly collimated beam of singly-charged and of multiply-charged particles; most of the multiply-charged particles are relativistic α-particles. These showers containing strongly collimated relativistic α-particles are characteristic of stars produced by heavy nuclei and are not observed in proton or helium produced stars. A typical event is illustrated in Fig. 4.

Perkins[7] and others have shown that α-particles ejected in the evaporation process are emitted isotropically with an average energy of 2–3 Mev per nucleon and that very rarely α-particles are emitted with more than 7.5-Mev/nucleon.

It follows that the energy and angular distribution of these α-particles in the laboratory system permit a measurement of the velocity with which the evaporating nucleus moved with respect to the laboratory frame of reference. We have obtained the energy of the incident nucleus both by measuring the angular spread of the α-particle showers and by measuring the multiple scattering of these α-particles.

(3) Energy Determination from the Opening Angle of the α-Particle Showers

For α-particles which are emitted isotropically in the rest system of the incident nucleus, the root-mean-square angle which the shower particles make in the laboratory system with the direction of motion is simply related to their mean kinetic energy in the rest system:

$$\langle \theta^2 \rangle = \langle T \rangle M / 3 p^2. \qquad (1)$$

Here $\langle T \rangle$ is the average kinetic energy of the α-particles in the rest system, M is the proton mass, and p is the momentum per nucleon of the incident particle.

Perkins[7] gives the average kinetic energy of evapora-

[7] D. H. Perkins, Phil. Mag. 41, 138 (1950).

tion α-particles as 8 Mev in carbon and oxygen and as 14 Mev in silver and bromine. Since most of the particles whose energy was measured by this method belonged to the magnesium-silicon group we have chosen for $\langle T \rangle$ the value of 10 Mev. One obtains then for incident relativistic particles

$$\langle \theta^2 \rangle^{\frac{1}{2}} = 0.056/E, \tag{2}$$

where E is the total energy per nucleon of the incident particle in Bev. As can be seen from Eq. (1) this result is insensitive to the accurate value of $\langle T \rangle$.

If the shower contains very few α-particles, the energy estimate based on the opening angle could be quite wrong, since in the rest system all the α-particles may accidentally be emitted in the forward or backward direction or because they may all have energies well below the assumed average value of 10 Mev. Both effects would decrease the average opening angle and lead to an overestimate of the incident energy. However, as the multiplicity of the α-particles increases, the danger of a gross overestimate decreases rapidly.

If we assume that none of the α-particles can be emitted in the rest system with kinetic energies larger than 30 Mev, the largest observed angle of any shower particle will provide an upper limit for the energy of the incident particle. The largest possible angle in the laboratory system is given by

$$E \tan\theta_{max} = \beta_{max}(\beta^2 - \beta_{max}^2)^{-\frac{1}{2}}$$

or

$$\theta_{max} \approx 0.12/\mathrm{E}. \tag{3}$$

Since $\theta_{max} \geqslant \langle \theta^2 \rangle^{\frac{1}{2}}$ it follows that the incident energy cannot be underestimated by more than a factor 2.

In order to obtain the opening angle of the α-particle showers, the following procedure was used:

(a) An area of 155 cm² of the plate on top of the "White Sands stack" was surveyed for all particles of atomic number $Z \geqslant 10$.

(b) The nuclei were identified by measuring the δ-ray density in the G-5 emulsions and the grain density in the under-developed C-2 emulsions which were placed on top of the brass stack. (See reference 3.)

(c) The tracks were followed through the brass stack until they left the stack or suffered a nuclear collision.

Of the 678 tracks found in this survey, 286 traversed the stack, 145 suffered collisions in the brass without leaving any multiply-charged particles proceeding in the original direction (A-collisions), 241 suffered collisions in which some multiply-charged fragments proceeded in the direction of the incoming track (B-collisions), and 6 stopped by ionization (see Chapter IV).

Of the 241 particles which suffered B-collisions, 77 gave rise to a shower containing at least 3 multiply-charged fast particles (mostly α-particles). We shall call these collisions "C-collisions." The energy measurements were carried out with this group. The shower was followed through additional brass plates and the position of the shower tracks was measured in successive plates.

Whenever the shower contained a fragment of charge $Z \geqslant 6$, its trajectory was assumed to define the direction of motion of the incident particle. In this case the angles

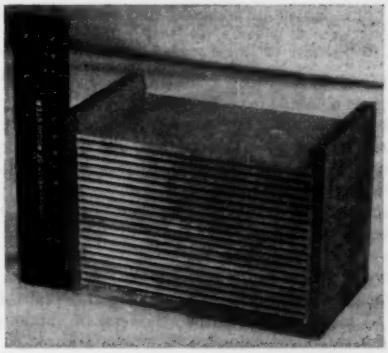


FIG. 3. Brass box used for exposing plates.

θ_i with respect to the direction of motion could be measured and the incident energy determined from Eq. (1). In most cases, however, the shower contained only α-particles and the angles ϕ_i of individual tracks were measured with respect to the center of gravity of the shower. The relation between the average value

$$\langle \phi^2 \rangle = \sum_1^n \frac{\phi_i^2}{n}$$

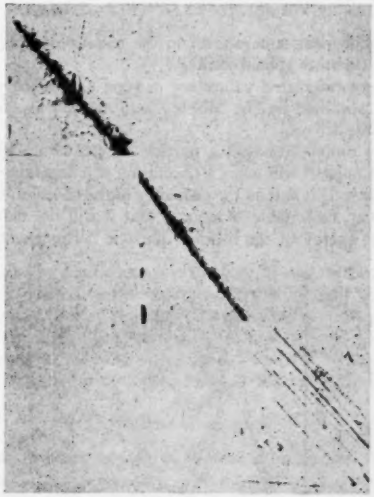


FIG. 4. C-collision: A primary Fe nucleus with a total energy of 2×10^{12} ev strikes a nucleus of the emulsion and breaks up into a highly collimated shower of 6 α-particles, a Be nucleus, and several protons. The first plate shows the incident nucleus before breakup. The second plate (4.6 mm from the first) shows the collimated shower shortly after breakup. The third plate (6.8 mm from the second) shows the separate components of the shower. The total spread in plate 3 is 30μ.

and $\langle\theta^2\rangle$ is derived in Appendix I(A) and is given by

$$\langle\phi^2\rangle=[(n-1)/n]\langle\theta^2\rangle, \tag{4}$$

where n is the number of tracks in the shower. Again the energy is obtained from Eq. (1).

Because of the finite number of particles in each shower, the energy values obtained by this method will fluctuate around the true value. By means of the Monte Carlo method we have determined the probability of obtaining a given value $\langle\phi^2\rangle$ for the average square angle of the shower particles with the center of gravity of the shower and thereby obtaining a given value ϵ' for the energy of the primary nucleus, if the true energy of the nucleus is ϵ. In Fig. 5 we have plotted the probability $P(\epsilon'/\epsilon)$ for (3- and 4-particle showers) of obtaining a value ϵ' if the true value is ϵ. The curve is based on the energy spectrum of evaporation α-particles given by Perkins.

From the mean energy $\langle T\rangle$ and the mean square energy $\langle T^2\rangle$ of Perkins' spectrum we can also calculate the mean square angle between the shower particles and the center of gravity,

$$\langle Y\rangle=\sum_i \frac{\phi_i^2}{n}=\frac{n-1}{3n}\frac{\langle T\rangle M}{p^2} \tag{5}$$

and the mean square deviation of Y:

$$\Delta^2=\langle(Y-\langle Y\rangle)^2/(\langle Y\rangle)^2\rangle=\frac{6}{5n}\frac{\langle T^2\rangle}{(\langle T\rangle)^2}-\frac{2n-3}{2n(n-1)}. \tag{6}$$

The root-mean-square error in the measured energy is approximately given by $\Delta/2$.

These calculated values are in good agreement with those obtained by the Monte Carlo method as shown in Table I.

The probability that a particle incident on a given target nucleus will make a C-collision will depend on its atomic weight and on the geometry of the collision. It is, however, reasonable to assume that it will not depend on the energy of the incident particle if this energy is

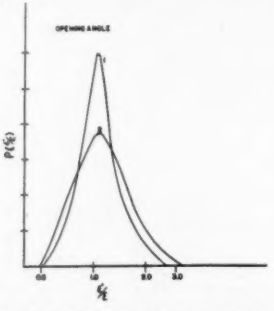


FIG. 5. Probability (in arbitrary linear units) of obtaining an energy ϵ' if the true value is ϵ for the opening angle measurements. Curves 1 and 2 are normalized to equal areas and refer repsectively to 4- and 3-particle showers.

TABLE I. Root-mean-square error, $\Delta/2$, in the measured energy.

	$\frac{\langle Y\rangle_{\text{Calc.}}}{\langle Y\rangle_{\text{M.C.}}}$	$\Delta/2$ Calc.	$\Delta/2$ M.C.
3-particle showers	0.99	0.25	0.29
4-particle showers	1.00	0.21	0.23

very large compared to the nuclear binding energy. This follows from the model which attributes the α-particle shower to the subsequent evaporation of particles in the cooling process of the nuclear matter which has escaped disintegration in the collision. If this assumption is granted, the particles which suffer C-collisions represent a sample of the incident particles randomly selected with respect to their energy, and the energy distribution obtained represents, after correcting for the ionization loss in the residual atmosphere, the energy distribution of the incident radiation.

If the angle between the shower particles is large, the particles may already be well separated in the emulsion following the brass plate where the collision occurs, and some of the shower particles could be missed. The probability of detecting a C-collision is, however, very nearly unity if we confine ourselves to small opening angles corresponding to energies above 3.0 Bev per nucleon. This value therefore represents the lower limit for which the measurements described here can be assumed to represent the primary energy spectrum.

The upper limit for the validity of this method is determined by the smallest opening angle which can be measured. Angles of the order of 10^{-5} radian can easily be measured by measuring the separation of tracks in emulsions separated by 10 cm of brass. As shown in the next section the relative scattering of the α-particles over this distance is not large enough to affect the opening angle measurement. The upper limit therefore corresponds to an energy of $\sim$5000 Bev/nucleon.

The practical limit is slightly lower and is determined by the destruction of α-particles by nuclear collisions in the brass with a mean-free path of 9 cm. The highest energy particle actually observed in this experiment was a silicon nucleus with 45 Bev/nucleon or a total kinetic energy of 1.3×10^{12} ev.

In Fig. 6 we have plotted the number of particles whose energy exceeded ϵ Bev/nucleon *versus* ϵ. The two curves in the same graph represent the values which we would expect to find if we take into account the probability distribution of Fig. 5 and assume for the incident radiation the energy spectra indicated in the graph. Because of the reduced detection probability of C-collisions produced by low energy particles, the points below $\epsilon=3$ Bev/nucleon do not represent the true number of incident primaries and should be disregarded. Figure 6 shows that the integral energy spectrum for particles of charge $10\leqslant Z\leqslant 26$ in the energy region between $3\leqslant\epsilon\leqslant 20$ Bev/nucleon is therefore well represented by

$$N(\epsilon)=K/(1+\epsilon)^{1.55\pm0.15}. \tag{7}$$

We shall show later that the same spectrum fits the data for all primaries of charge $2 \leqslant Z \leqslant 26$ also in the region of nonrelativistic energies.

(4) Energy Determinations from the Relative Scattering of α-Particles Produced in C-Collisions

Because the α-particles forming the shower characterizing a C-collision are produced in the rest system with kinetic energies below 30 Mev or 7.5 Mev/nucleon, their total energy per nucleon in the laboratory system E_0 will very nearly be equal to that of the parent nucleus (E) if the latter is relativistic. If ϑ is the angle in the rest system between the direction of motion of the parent nucleus and the emitted α-particle, the total energy per nucleon of the α-particle in the laboratory system in units of the proton mass is given by

$$E_0 = E_0'E + [(E_0')^2 - 1]^{\frac{1}{2}}(E^2-1)^{\frac{1}{2}} \cos\vartheta. \quad (8)$$

For $E_0' < 1.008$ (corresponding to the upper limit of 30 Mev for the α-particle energy in the rest system) we get

$$(|E_0 - E|)/E \leqslant 13 \text{ percent}. \quad (9)$$

We therefore obtain a nearly monoenergetic beam of α-particles whose average enery per nucleon is equal to that of the incident particle. The energy of the incident nucleus can therefore be determined by measuring the energy of the α-particles in the laboratory system.

The determination of particle energies from multiple coulomb scattering in emulsion has been described by several authors.[8,9] The method as previously applied is limited to energies below a few Bev per nucleon. In order to make measurements in the energy range between 1 and 100 Bev/nucleon it is necessary to reduce the reading error by using larger cell lengths than those available for tracks in a single emulsion and to avoid errors arising from the distortion of the emulsion and from inaccuracies of the microscopic stage.

The first can be accomplished in our arrangement by measuring the position of the track in different emulsions separated by several millimeters of absorber and by tracing the shower through distances of 10–20 cm.

Inaccuracies of the stage motion are eliminated by measuring only the relative scattering of the particles, i.e., measuring the relative separation of the shower tracks in successive plates.

The effect of distortions in the emulsion can be avoided by measuring the distances between tracks at the point where they enter the glass of the plate. The accurate location of the entrance point was facilitated in our experiment:

(a) by using emulsions of high sensitivity (the grain density for a relativistic α-particle was >130 g/100μ).

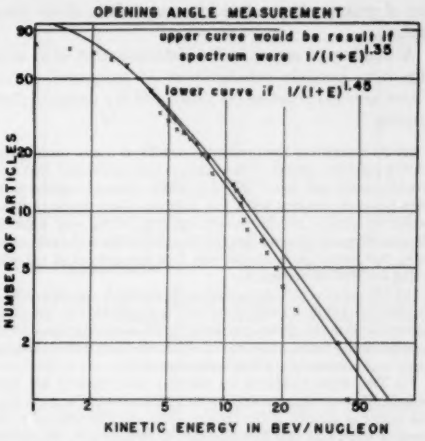


FIG. 6. Integral energy spectrum obtained from opening angle measurements.

(b) by using thin emulsions (100μ) which will not peel even if the usual coating between emulsion and glass is omitted,

(c) by using plates in a horizontal position such that most tracks are steeply dipping.

If X, X', X'' are the x or y coordinates of a given track at successive intervals of length t, the scattering S is defined by $S = X + X'' - 2X'$ and the energy is related to the average value of S^2 by

$$(\langle S^2 \rangle)^{\frac{1}{2}} = (Kt^{\frac{1}{2}})/p\beta, \quad (10)$$

where p, β are the momentum per nucleon and the velocity of the α-particle. For scattering in emulsions the constant K has been evaluated by Gottstein *et al.*[10] and agrees with the value calculated from a formula given by Williams.[11]

The value of K for the mixture of brass, glass, and emulsion used in our experiment is calculated in Appendix II.

Since in our experiment the coordinates of the track are measured not with respect to a straight line of reference, but with respect to the line connecting the centers of gravity of the shower in successive plates, our measurements yield a quantity $\langle S^2 \rangle$ which is related to the true scattering by the relation derived in Appendix I(B),

$$(\langle S^2 \rangle)^{\frac{1}{2}} = [(n-1)/n]^{\frac{1}{2}}[\langle (S')^2 \rangle]^{\frac{1}{2}},$$

where n is the number of tracks in the shower. In our method of determining the scattering S, we have eliminated errors which arise from inaccuracies of the microscope stage and distortion of the emulsion. Our main error comes in determining the point where the track enters the plane of the glass, because of the finite

[8] P. Fowler, Phil. Mag. 41, 169 (1950).
[9] Y. Goldschmidt-Clermont, Nuovo cimento 7, 1 (1950).
[10] K. Gottstein *et al.*, Phil. Mag. 42, 708 (1951).
[11] E. J. Williams, Phys. Rev. 58, 292 (1940).

size of grains and possible inhomogeneities of the glass surface.

A number of errors in the determination of S arise also from possible misalignment of successive plates. These are: (a) rotation, (b) tilt, and (c) unequal plate spacing.

(a) By comparing the position of tracks of heavy nuclei penetrating opposite corners of the stack, it was established that relative rotation is less than 1/300 of a radian. Since no measurements were made on showers when the distance of any track from the center of gravity of the shower exceeded 600μ, and since for showers of energy above 5 Bev/nucleon the distance did not exceed 300μ, this correction amounts only to a few percent of the scattering and can be neglected.

(b) The relative tilt of successive plates which was measured to be less than 1/500 of a radian has still less influence on the measurements since the errors produced by tilt are proportional to the tangent of the zenith angle under which the track enters, and this angle was almost always much less than 45°.

(c) The errors produced by unequal plate spacing are more serious. They produce variations of the distance of the track from the center of gravity of the shower which are proportional to the opening angle and therefore inversely proportional to the particle energy. Since the scattering is also inversely proportional to the energy, the spurious scattering due to variation of plate spacing produces a percentage error in the scattering measurements which is the same for showers of any energy. The variation of individual plate spacing amounted to about 0.1 mm or 2 percent of the average plate spacing of 5 mm. Over the distance of one plate spacing the spurious scattering amounts therefore to about 40 percent of the true scattering. Since this number is independent of the energy of the shower and since errors must be added in a Gaussian manner, the error introduced is not too serious. Fortunately, however, it can be eliminated. The spurious scattering produced by errors in plate spacing increases or decreases the separation of tracks from the center of gravity by the same percentage, while the true scattering changes these distances in a random manner. Consequently the error introduced by variation in plate spacing can be eliminated in the statistical treatment of the data. This is done in Appendix IC.

Our total spurious scattering resulting from noise was determined to be $\sim 1\mu$.

The probability $P(\epsilon'/\epsilon)$ of obtaining from the scattering measurements an energy value ϵ' if the true energy is ϵ, has been calculated on the basis of an assumed gaussian distribution in S, leading to a Γ-distribution in $\langle S^2 \rangle$. The resulting probability distribution is shown in Fig. 7. The results of the scattering measurements have been plotted in Fig. 8, where they can be compared with the curves which give the results to be expected from these measurements if the true energy distribution of the particles at the top of the atmosphere followed the distribution laws indicated in the graph.

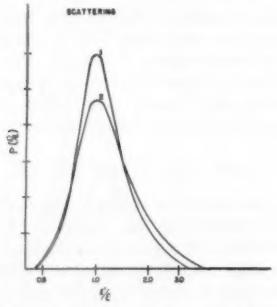


FIG. 7. Probability (in arbitrary linear units) of obtaining an energy ϵ' if the true value is ϵ for the scattering measurements. Curves 1 and 2 are normalized to equal areas and refer, respectively, to 6 and 4 independent readings.

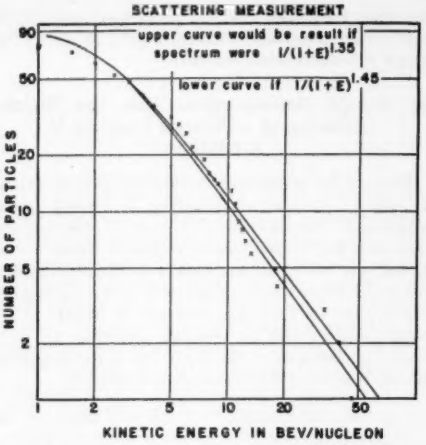


FIG. 8. Integral energy spectrum obtained from scattering measurements.

The two independent methods by which the energy of particles producing C-collisions is obtained not only give the same energy spectrum, they also give particle energies which in almost all cases are in good agreement with each other. We have calculated on the basis of the probability distributions shown in Figs. 5 and 7 the expected distribution in the ratio of energy values determined by the two methods. Figure 9 shows the expected and observed ratios and indicates that the results obtained by the two methods are consistent with each other.

Table II lists a random sample of events:

(a) The first column identifies the primary nucleus.

(b) The second column lists the particles of charge $Z>1$ in the shower.

(c) The third column gives the cell length used for the scattering measurements.

(d) The fourth column gives the number of independent scattering measurements made. This number is $(2n-3)(c-1)$ where c is the number of cells and n the number of shower particles. The number 2 is due to the fact that scattering is measured in two mutually perpendicular directions. The factor $c-1$ appears because 2 cells are required to establish second differences. The number 3 in the bracket arises because in each measurement two readings are required to refer the scattering to the center of gravity of the shower and one reading to correct for errors in plate spacing.

(e) The fifth and sixth columns give the energy values obtained from scattering and mean square opening angle, respectively.

(f) The last column is an upper limit on the energy of the particle

derived from the largest angle in the shower as explained in Chapter I, (3).

The method of measuring the energy of primary particles outlined in this chapter is limited to particles of charge $Z \geqslant 10$ because only very heavy nuclei produce showers with 3 or more α-particles with appreciable probability. The interpretation of the results in terms of a primary energy spectrum is further dependent on the assumption discussed in I, (3), that the probability of C-collisions is independent of the particle energy at high energies. This assumption, while very reasonable, cannot as yet be proved; however, even a variation of this probability of as high as 20 percent in the energy range from 3–10-Bev/nucleon would not affect our results within the statistical error. In Chapter V it is shown that the interaction cross section of heavy nuclei does not strongly depend on energy.

It may be possible to overcome these limitations at least partially by using absorbers rich in hydrogen instead of brass between the emulsions. The probability that a collision between a very heavy nucleus and a proton results in a measurable shower may then approach unity and may be appreciable even for primaries of the carbon, nitrogen, oxygen group.

Another possible method of extending these techniques, however, consists in measuring the scattering of single unrelated tracks with respect to the track of a particle of very high energy. In Chapter VI it is shown that particles with energies between 10^{11} and 10^{12} ev can easily be found and identified in our brass stack and can therefore be used as a reference line for scattering measurements on any other penetrating particle.

For the purpose of carrying out measurements described in this section, Bausch and Lomb Optical Company have constructed a special microscope according to our specifications. This microscope possesses an accurate screw by which the plate, rotated through an arbitrary angle, can be moved over long distances which are easily measurable with an accuracy of 0.2μ.

CHAPTER II. THE NONRELATIVISTIC PART OF THE SPECTRUM FOR PRIMARY COSMIC-RAY PARTICLES OF CHARGE $Z \geqslant 6$

The energy of heavy nuclei below 1 Bev/nucleon can be determined by measuring their range in an absorber; these nuclei have an appreciable chance of slowing down and coming to rest because of the high ionization loss of multiply charged particles. Tracks of particles which are slowing down show an appreciable increase in δ-ray density or grain density along their trajectory, and can be identified by methods previously described.[1,2] For determining the range and thereby the energy spectrum of heavy nuclei we have used the brass stack flown at geomagnetic latitude $\lambda = 55°$ described in Chapter I. This stack is well suited for such measurements because a particle entering from the vertical has to traverse 60 g/cm² of brass and glass before leaving the stack and its

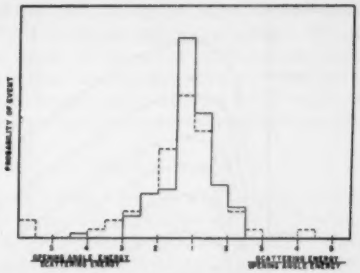


FIG. 9. Probability (in arbitrary linear units) of obtaining a given ratio for the energies as determined by opening angle and scattering measurements (the larger energy has been taken as the numerator). The dotted lines indicate the observed results and are normalized to the same area as the calculated values.

δ-ray density can be observed at intervals of 5 mm corresponding to 3 g/cm².

By limiting our investigation to those tracks which entered in a direction presenting at least 60 g/cm² of absorbing material to the particle, we did not bias our selection as far as particle energy is concerned because at that latitude the primary flux is known to be isotropic in zenith angle.[2]

At the latitude of Minnesota the earth's magnetic field permits the entry of particles of $Z > 1$ if their energy exceeds 0.3 to 0.35 Bev per nucleon. However at the altitude at which our experiment was carried out ($\sim$20 g/cm² of residual atmosphere), particles whose energy was close to the cut-off value could reach the stack only if their atomic number was below that of neon. Our experiments therefore yield an energy spectrum for the CNO group for energies $\epsilon > 0.35$ Bev/nucleon and for heavier particles for energies $\epsilon \gtrsim 0.55$ Bev/nucleon.

Particles found in a survey of the top emulsion were followed through the stack until they either had

TABLE II. Energy of heavy primary nuclei determined from their breakup products.[a]

Primary	Shower particles	Cell length mm	Number of independent readings	Energy from scattering Bev/nuc.	Energy from opening angle Bev/nuc.	Upper limit of energy from opening angle Bev/nuc.
Fe^{26}	$Be^4+6\alpha$	27.0	17	32.5	23.0	37.0
Na^{11}	4α	13.7	5	4.7	4.8	10.0
Mg^{12}	4α	5.2	5	3.9	3.6	8.0
Ca^{20}	$B^8+4\alpha$	6.0	7	1.95	1.9	3.5
Si^{14}	$Ne^{10}+2\alpha$	5.1	9	1.3	2.1	4.5
Si^{14}	$Li^7+3\alpha$	11.1	8	11.1	11.2	22.0
Al^{13}	3α	10.3	12	7.4	7.8	16.8
Mg^{12}	4α	6.5	5	1.5	1.5	2.8
Al^{13}	4α	5.4	5	0.64	0.56	1.5
Si^{14}	4α	16.0	3	46.0	44.0	75.0
Si^{14}	$F^9+2\alpha$	6.2	9	5.0	3.7	6.8
Fe^{26}	$O^8+3\alpha$	6.4	5	0.36	0.69	1.7

[a] The range of the O^8 fragment in the last shower indicates the energy lies between 0.37 and 0.42 Bev per nucleon.

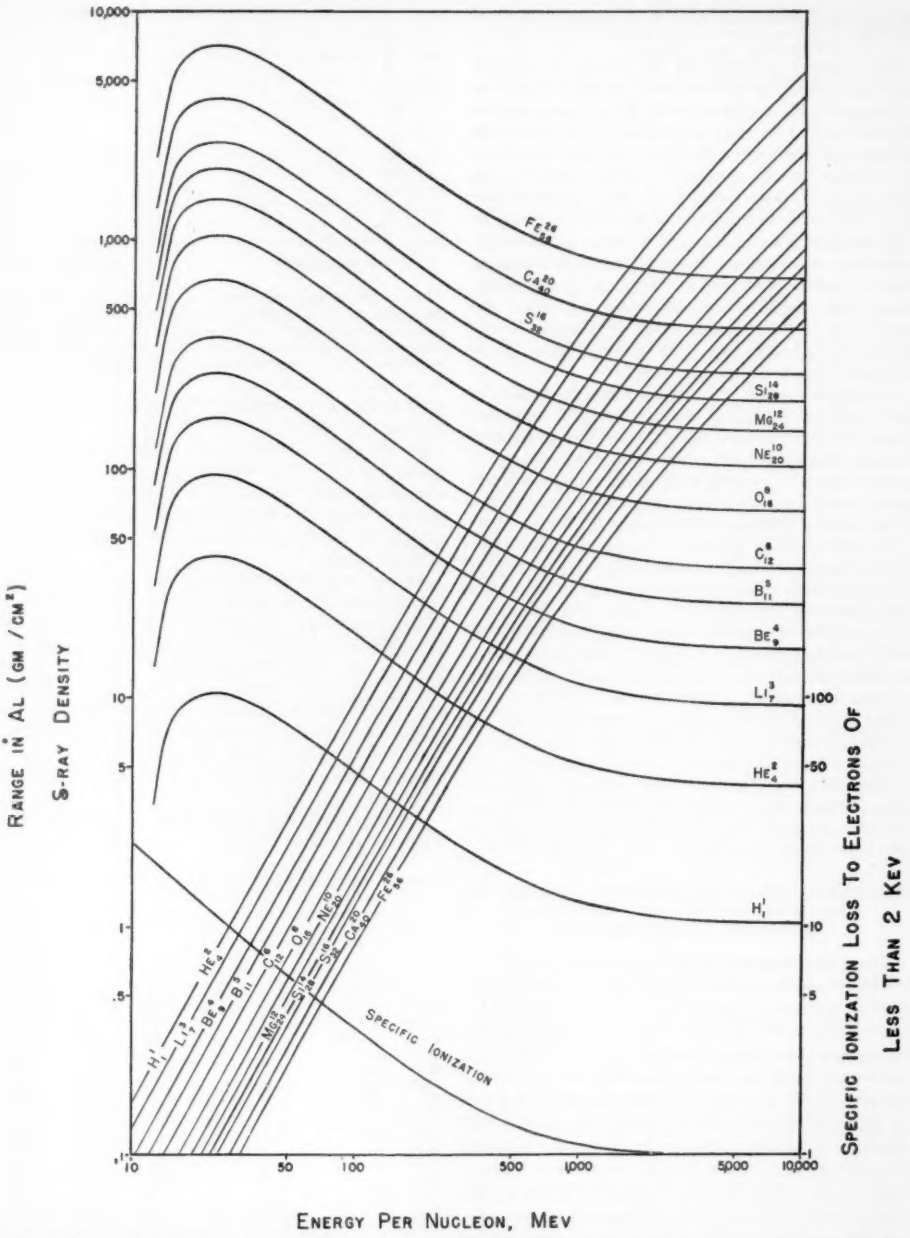


FIG. 10. Range, δ-ray density, and specific ionization loss to electrons of less than two kev as a function of energy in Bev per nucleon for various nuclei.

traversed 60 g/cm² or had disappeared in the brass plate between two emulsions. If the track showed a consistent increase in δ-ray density in the plates preceeding its termination it was classed as a particle of an ionization range determined by the length of its trajectory from the top of the atmosphere to its end point. If the particle did not change its δ-ray density it was classed as a particle which had suffered a nuclear collision.

The uncertainty in the measured range of stopping particles is of the order of the thickness of the brass absorbers (3 g/cm²) and, being only a small fraction of the total range from the top of the atmosphere, produces only a small error in the range spectrum.

The range-energy relation for particles of different charge and mass are known since they are simply derived from the range-energy relation for protons

$$R(Z, M, \epsilon) = (M/Z^2)R(1, 1, \epsilon),$$

where ϵ is the energy per nucleon. Figure 10 shows the dependence of R on ϵ for particles of different atomic numbers. The greatest uncertainty in converting our range spectrum into an energy spectrum arises from errors in the charge determination. The charge was determined with the help of the sensitive and insensitive emulsions (see Appendix III) on the top of the stack by methods previously described.[3,12] However, since in this experiment horizontal plates were used and therefore the tracks investigated were steeply dipping and short, the error in charge determination is greater than in previous experiments and probably as large as 20–25 percent. This fact and the limited number of tracks investigated restrict us to establishing an energy spectrum for two groups of primary nuclei:

$$6 \leqslant Z \leqslant 9 \quad \text{and} \quad Z \geqslant 10.$$

In reducing the observed range spectrum to an energy spectrum of particles at the top of the atmosphere, we also have to make a correction for those particles which suffered nuclear collisions before they came close to the end of their range. This correction was made on the assumption that for particles of energy $0.35 < \epsilon < 1.5$-Bev/nucleon the cross section for nuclear collision is independent of energy. However, even if the nuclear cross section should vary by as much as 50 percent in this energy interval, the spectrum obtained from the range data would hardly be affected.

An area of 12.5 cm² was surveyed for primaries of the CNO group. We found a total of 136 particles of which 33 traversed the stack, 70 suffered nuclear collisions, and 33 came to the end of their range. A similar survey in an area of 29 cm² for particles of charge $Z \geqslant 10$ yielded 77 particles of which 12 traversed the stack, 43 suffered nuclear collisions, and 22 came to the end of their range. It was found that of the particles whose range exceeded R g/cm² the fraction of particles coming to rest in the interval between R and $(R+10)$ g/cm² was: 7.8±1.4 percent for the CNO group and 9.4±2.4 percent for the $Z \geqslant 10$ group in the range interval $27 < R < 80$ g/cm². Figure 11 shows the integral energy spectrum obtained from these measurements, as well as flux values obtained at geomagnetic latitudes $\lambda = 55°$ and $\lambda = 41.7°$.

The low energy data can be represented by an integral spectrum:

$$N(\epsilon) = K/(\alpha+\epsilon)^{1.25},$$

where $\alpha = 0.9 \pm 0.2$ for the CNO group, $\alpha = 1.0 \pm 0.3$ for $Z \geqslant 10$, and ϵ is the kinetic energy in Bev/nucleon. This is in agreement with the empirical spectrum derived for the high energy region.

CHAPTER III. THE DETERMINATION OF THE ABSOLUTE INTENSITY OF DIFFERENT COMPONENTS OF PRIMARY RADIATION

The survey of the plates on top of the brass stacks flown at geomagnetic latitude $\lambda = 41.7°$ and $\lambda = 55°$ can also be used to obtain the flux of heavy nuclei at these latitudes. The necessary corrections for nuclear absorption have been described previously.[3] A charge calibration for the tracks was achieved by comparing their δ-day density with that of α-particles in the sensitive emulsions and their grain density with that of slow α-particles in the insensitive emulsions (see reference 3 and Appendix III). At $\lambda = 41.7°$ (cut-off energy 1.5 Bev per nucleon) the particles enter the stack with relativistic velocity. The δ-ray density for relativistic carbon tracks was found to be 3.7 δ-rays per 100μ.

(1) Flux at Geomagnetic Latitude $\lambda = 41.7°$

The survey of the top emulsion was carried out in such a way as to include all particles whose δ-ray density exceeded 2.5 δ-rays/100μ. In 12.25 cm² we

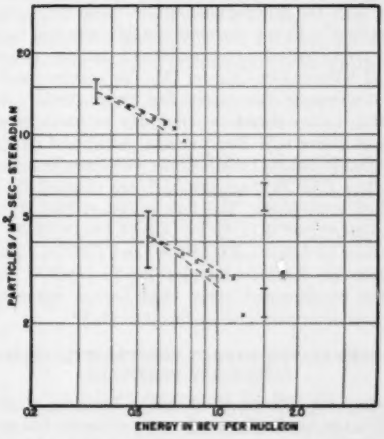


FIG. 11. Integral energy spectrum obtained from range measurement. The dotted lines refer to the errors quoted in the text.

[12] H. Bradt and B. Peters, Phys. Rev. 74, 1828 (1948).

found 100 particles which penetrated at least 8 g/cm² of absorber. This survey should therefore include all particles of charge $Z \geqslant 6$. In order to test this point an additional area of 1 cm² was surveyed to include all tracks whose δ-ray density exceeded 1.7 δ-rays per 100μ. Fifty-one additional tracks were recorded but none of them penetrated 8 g/cm² of absorber, proving that these tracks were caused by slow secondary singly or doubly charged particles.

It also suggests that primary nuclei of beryllium or boron with an energy above 1.5 Bev/nucleon are rare or absent at this latitude. This is in agreement with our previous result at $\lambda = 30°$ [3] but in disagreement with the result obtained by Dainton *et al.*[13] at $\lambda = 54°$.

The flux of primaries at $\lambda = 41.7°$ was determined as $I(41.7°) = 5.9 \pm 0.7$ particles/m² sec steradian for $6 \leqslant Z \leqslant 9$ and $I(41.7°) = 2.4 \pm 0.3$ particles/m² sec steradian for $Z \geqslant 10$.

The errors given are the statistical standard deviation. The upper value is in very good agreement with results recently obtained by Van Allen[14] at the same latitude with an ionization chamber mounted on a rocket. The value for particles of charge $Z \geqslant 10$ is twice as high as Van Allen's value.

(2) The Flux at Geomagnetic Latitude $\lambda = 55°$

In order to obtain from the number of tracks observed to enter the stack at $\lambda = 55°$ the flux of particles at the top of the atmosphere, it is necessary to correct not only for particles lost by nuclear collisions in the residual atmosphere, but also for those particles which when arriving at large zenith angles have not sufficient range to reach the stack. Since the range spectrum has been determined (Chapter II), this correction can be made down to energies determined by the atmospheric cutoff for vertical incidence. This atmospheric cut-off value is lower than the geomagnetic cut-off value for particles of charge $Z < 10$ and therefore the flux obtained for the CNO group includes presumably all particles entering at that latitude (see Chapter IV). For heavier particles the atmospheric lies above the geomagnetic cutoff; our flux values therefore refer only to particles with energies above 0.55 Bev/nucleon. Additional particles presumably arrive at this latitude with energies between 0.35 and 0.55 Bev/nucleon and are stopped by the residual atmosphere. The flux values at this latitude were determined as: $I(55°) = 14.4 \pm 1.4$ particles/m² sec steradian for CNO ($\epsilon \geqslant 0.35$ Bev) and $I(55°) = 4.2 \pm 1.0$ particles/m² sec steradian for $Z > 10$ ($\epsilon \geqslant 0.55$ Bev). Within experimental error these values agree with previous determinations at this latitude.[3, 4]

CHAPTER IV. THE ENERGY SPECTRUM OF PRIMARY COSMIC-RAY PARTICLES

Before the latitude effect, which has been observed both in the total radiation (mostly protons) and in the heavy component, can be interpreted in terms of an energy spectrum, it is necessary to consider two problems:

(a) The determination of the number of primary protons entering the atmosphere is complicated by the fact that they cannot easily be distinguished from secondary high energy protons produced in collisions in the upper atmosphere. This difficulty is not eliminated if the flux is measured outside the atmosphere by telescopes carried in rockets. Secondary protons are leaving the atmosphere and may even return again after being bent back by the earth's magnetic field at large distances. That this effect is appreciable was shown by Van Allen,[15] who obtained at a height of about 100 km a counting rate which increased strongly with zenith angle. The observed angular distribution $I(\theta°) = I(0°)(1 + 0.6 \sin\theta)$ was interpreted as due to an appreciable atmospheric albedo, which may be expected to increase with zenith angle. The values for the total primary flux collected by Winckler *et al.*[5] represent therefore still an upper limit to the true flux.

Apart from its effect on the absolute flux value, the albedo may also produce a distortion of the energy spectrum since it may be expected that the more energetic primaries give a larger contribution to the albedo than primaries of low energy. In this case the total counting rate at the top of the atmosphere will fall off more slowly with latitude than the primary intensity. For this reason, the exponent in the energy spectrum given by Winckler *et al.*, which refers mainly to the proton component and can be represented by $N(\epsilon) = K/(1+\epsilon)^{1.07}$ may be too low.

(b) The difficulty which arises in the determination of the primary proton flux does not arise in determining the flux of heavy primaries. Collisions in the upper atmosphere will either destroy the particle completely or result in some heavy fragments which will proceed undeviated in the original direction and cannot leave the atmosphere.

However, in trying to convert the latitude effect for heavy primaries into an energy spectrum we encounter another problem: we must know the charge to mass ratio of these nuclei at the time they enter the influence of the earth's magnetic field, i.e., far outside the atmosphere.

We propose to show that even the heaviest nuclei observed in the primary beam ($Z \sim 26$) are completely stripped of electrons when they come under the influence of the earth's magnetic field, and that therefore their charge to mass ratio is known to have the value $Z/M \approx \frac{1}{2}$.

Some indication of this can already be obtained from the fact that the direct energy measurements described in Chapters I and II agree with the spectrum which one obtains from the latitude effect if one assumes that the nuclei arrive completely stripped of electrons.

[13] A. D. Dainton *et al.*, Phil. Mag. **42**, 317 (1951).
[14] J. A. Van Allen, private communication.
[15] J. A. Van Allen and A. V. Gangnes, Phys. Rev. **78**, 50 (1950).

TABLE III. Cut-off data for particles at geomagnetic latitude $\lambda=41.7°$.

	Cut-off energy in Bev/nuc. completely stripped	Range in g/cm² Al	Cut-off energy in Bev/nuc. with K-shell present	Range in g/cm² Al	Cut-off energy in Bev/nuc. with K and L shell present	Range in g/cm² Al
Fe	1.35	49.6	1.21	43.0	0.655	18.2
Ca	1.50	69.0	1.29	58.0	0.525	16.0
Si	1.50	98.6	1.21	65.6	0.20	4.7
Mg	1.50	115.	1.16	83.4	0.07	0.87
Ne	1.50	138.	1.09	92.0	0	0

An investigation whether or not very heavy nuclei appear in the cosmic radiation with energies below the geomagnetic cutoff calculated for particles of $Z/M=\frac{1}{2}$ cannot easily be carried out at northern latitudes because it requires exposure of plates at altitudes not yet accessible to balloons. At $\lambda=41.7°$, however, such an investigation is possible.

We have two identical brass stacks which were flown under nearly identical conditions at $\lambda=55°$ and $\lambda=41.7°$. Since as described in Chapter II we know the intensity of the low energy component at $\lambda=55°$, we can calculate how much of this component would have been excluded at $\lambda=41.7°$ if the particles entered the earth's field: (α) completely stripped of electrons, (β) retaining only the K shell, and (γ) retaining both the K and L shells. From this we can calculate how many of the particles should have come to the end of their range and stopped in the White Sands stack under each of the foregoing assumptions and compare it with the actual number of stopping tracks observed.

Table III shows the cut-off energy and cut-off range for different nuclei at $\lambda=41.7°$ under the three different assumptions given above. Table IV shows the number and type of nuclei which came to the end of their range in the White Sands stack as well as the number expected under the three different assumptions. As can be seen from Table IV the heavy nuclei arrive either completely stripped or retain at most one or two electrons.

If we make the assumption that heavy nuclei are only partially ionized at the source of cosmic radiation, as seems to be the case for all regions of the universe accessible to spectral analysis, we obtain a lower limit for the amount of matter which the particles must have traversed before approaching the earth. Making use of calculations given by Williams,[16] the cross section for the loss of an electron from the L shell of iron due to the passage of a proton or electron is

$$\sigma=4.7\times10^{-22}\text{ cm}^2\text{ for }\beta=1.$$

It rises to a maximum:

$$\sigma=2\times10^{-19}\text{ cm}^2\text{ for }\beta=4.8\times10^{-2}$$

and from there it drops rapidly as the velocities decrease further. At the most favorable velocity an iron ion must therefore traverse at least 12μ g/cm² of ionized hydrogen gas in order to lose most of its L electrons and more if the intervening material is not ionized. This according to Van de Hulst[17] represents nearly 200 times as much material as the particle would encounter if travelling directly from the solar corona to the earth. The direct passage of heavy primaries from the solar corona to the earth seems, therefore, to be excluded.

Having established that heavy nuclei are completely stripped of electrons before approaching the earth we can now use the flux data at different latitudes in conjunction with information obtained from nuclear collisions (Chapter I) and range distribution (Chapter II) to construct the integral energy spectrum of primary cosmic radiation. The spectrum is shown in Fig. 12. The vertical bars represent flux measurements[18] corrected to the top of the atmosphere on the assumption of an incident isotropic radiation.[2] The dots and crosses represent the results obtained from opening angles and relative scattering of particles in α-particle showers normalized to the flux measurements at 3 Bev. In the helium spectrum we have used the values obtained with counters by Ney[19] at $\lambda=55°$ and Perlow[20] at $\lambda=41°$ as well as our own earlier measurements at $\lambda=51°$ and 30°.[2,4]

The proton spectrum is obtained from the spectrum given by Winckler[5] after subtracting the contribution of the heavy component. The total number of nucleons incident on the top of the atmosphere as a function of energy is also shown in the graph. The curves drawn for the heavy nuclei represent the spectrum

$$N(\epsilon)=K/(1+\epsilon)^{1.65},$$

where ϵ is the kinetic energy per nucleon in Bev and K has the values $K=400$ particles/m² sec steradian for helium, $K=20$ particles/m² sec steradian for the CNO group, and $K=8$ particles/m² sec steradian for $Z\geq10$.

The latitude effect for primaries in the last two groups

TABLE IV. Particles stopping in stack.

	1 No. of particles which should be seen stopping if all are stripped	No. observed stopping	2 No. of particles which should be seen stopping if they had retained K-shells	3 No. of particles which should be seen stopping if they had retained K- and L-shells	No. observed below cut-off for completely stripped particles
Fe	3.8	3	5.4	17	0
Ca	1.5	1	2.8	13	0
Si	0.6	2	4.4	53	0
Mg	0	0	2.3	65	0
Total	5.9	6	14.9	148	0

[16] E. J. Williams, Kgl. Kanske Videnskab. Selskab, Mat.-fys. Medd. XIII, 4 (1935).

[17] H. C. Van de Hulst, Astrophys. J. 105, 471 (1947).

[18] The experimental point for the CNO group at $\lambda=0°$ represents preliminary unpublished data obtained from equatorial flights carried out in India by B. Peters *et al.* in October, 1950.

[19] E. P. Ney and D. M. Thon, Phys. Rev. 81, 1068 (1951).

[20] G. J. Perlow and L. R. Davis, private communication.

has also been measured by Ney *et al.*[21] Their results are in agreement with ours within the experimental error.

The energy distribution in Fig. 12 is, however, in disagreement with that reported by Dainton *et al.*,[13] who measured the energy spectrum of heavy nuclei incident at $\lambda=54°$ by multiple Coulomb scattering of tracks in single emulsions. Their latest data[22] suggest that only 28 percent of the particles incident at this latitude have energies above 1.5 Bev per nucleon and only 8 percent energies above 3.5 Bev/nucleon. The corresponding calues in our spectrum are 45 percent and 21 percent, respectively. It is difficult to see how the resuls of these authors can be reconciled with the observed intensity ratio of 4–5 between $\lambda=54°$ (cut-off energy 0.4 Bev/nucleon) and $\lambda=30°$ (cut-off energy 3.5 Bev/nucleon).

The proton spectrum in Fig. 12 can be represented by the formula

$$N(\epsilon)=4000/(1+\epsilon)^{1.07} \text{ particles/m}^2 \text{ sec steradian.}$$

As explained in the beginning of this section, the true exponent may be slightly higher and the numerator slightly lower, since the flux measurements include some secondary particles.

In any case, it seems that the velocity spectrum of all particles of charge $Z>1$ is independent of the mass of the particle and that the chemical composition of that part of the beam stays constant for energies between $0.35<\epsilon\lesssim 10$ Bev/nucleon.

A velocity spectrum independent of atomic number would be in agreement with the assumption that the particles are accelerated by some electromagnetic mechanism in which most of the energy is given to the particles after they have traversed several μ g/cm^2 of matter and attained identical charge to mass ratios. At the same time the amount of matter traversed before completing the acceleration should not be large enough to make nuclear collisions probable, because if the acceleration were terminated by a nuclear collision, the largest nuclei would spend the shortest time in the acceleration process and should therefore have a spectrum which falls off faster than that of smaller nuclei. Such assumptions could account for, but do not necessarily follow, from our results.

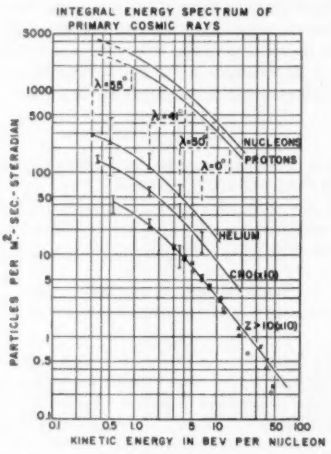


FIG. 12. Integral energy spectrum of primary cosmic radiation. The vertical bars represent flux measurements corrected to the top of the atmosphere on the assumption of an incident isotropic radiation. The dots and crosses represent the results obtained from opening angles and relative scattering measurements normalized to the flux measurements at 3 Bev. The proton spectrum is obtained from the spectrum of Winckler. The solid curves drawn for the heavy nuclei are represented by $N(\epsilon)=K/(1+\epsilon)^{1.35}$ and for the protons by $N(\epsilon)=4000/(1+\epsilon)^{1.07}$, where ϵ is kinetic energy in Bev per nucleon.

[21] P. Freier *et al.*, Phys. Rev. **84**, 384 (1951).
[22] A. D. Dainton *et al.*, private communication.

A velocity spectrum independent of atomic number would be in agreement with the assumption that the particles are accelerated by some electromagnetic mechanism in which most of the energy is given to the particles after they have traversed several μ g/cm^2 of matter and attained identical charge to mass ratios. At the same time the amount of matter traversed before completing the acceleration should not be large enough to make nuclear collisions probable, because if the acceleration were terminated by a nuclear collision, the largest nuclei would spend the shortest time in the acceleration process and should therefore have a spectrum which falls off faster than that of smaller nuclei. Such assumptions could account for, but do not necessarily follow, from our results.

CHAPTER V. NUCLEAR COLLISIONS OF HEAVY PRIMARIES

In the course of the work described in the previous chapters nearly 1000 particles were traced into the brass stack and their nuclear collisions were observed as described in Chapter II. In terms of g/cm^2 the stack was composed of 88 percent brass, 10.7 percent glass, and 1.3 percent emulsion. Since the mean free path of primaries in glass was determined previously,[3] the mean free path in brass can be obtained.

The collision mean free path was determined both in the Minnesota and White Sands stack. A total of 305 collisions was observed. The results for corresponding charge groups were the same within the statistical error. Since more than half of the particles in the Minnesota stack have energies below 1.5 Bev/nucleon, while all of the particles in the White Sands stack have energies above 1.5 Bev/nucleon, the results show that, as expected, the cross section for nuclear collisions does not depend strongly on particle energy.

A mean free path in brass has also been obtained for α-particles of kinetic energy above 10 Bev by tracing the shower particles described in Chapter I.

The results obtained in brass as well as our previous results in glass are given in Fig. 13. The experimental points are much higher that the dotted curves which represent the values predicted, on the assumption that the collision cross section is equal to the geometric cross section. The data for both materials are, however, well represented by the solid curves which correspond to the cross section:

$$\sigma=\pi(R_1+R_2-2\Delta R)^2, \qquad (11)$$

where the subscripts refer to the target and projectile,

$$R=1.45\times10^{-13}A^{\frac{1}{3}} \text{ cm}, \quad \Delta R=1.0\times10^{-13} \text{ cm}.$$

The validity of this relation, which had been suggested by previous data,[2] has now been tested for absorbers whose atomic weight lies between that of air and brass and for incident nuclei from helium to iron.

CHAPTER VI. FLUX AND ENERGY DETERMINATIONS IN THE RANGE OF $10^{11} \leq \varepsilon \leq 10^{14}$ EV

When scanning a plate well imbedded in the brass stack, one frequently encounters large numbers of parallel minimum ionization tracks, often grouped into several distinct cores. When following these showers in the downward direction the multiplicity of tracks increases rapidly. In most cases it reaches a maximum value and then decreases again. If followed into the upper hemisphere, one finally reaches the plate where the divergence of tracks permits constructing the location of the shower origin in the brass plate above. If this origin lies close to the emulsion, it is at times possible to observe the shower of the charged mesons before the γ-rays produced by neutral mesons have been converted. The latter are near the origin of the showers, recognizable by the close doublets which they produce.

Apart from giving information on the multiplicity, lifetime, energy distribution, and angular distribution of neutral mesons, on the multiplicity and angular distribution of charged mesons, and on the multiplication of the soft component, these showers permit the determination of the primary flux of protons and α-particles in the energy region between 10^{11} and 10^{13} ev.

We have at present 35 of such showers whose energy exceeds 10^{11} ev. Several of such events exceed 10^{12} ev and one exceeds 10^{13} ev.

The analysis of these events has not been completed. Of 9 showers of energy above 10^{11} whose primaries have been identified, four were probably produced by γ-rays entering from outside, four were produced in the stack by singly charged particles and one was produced by a primary α-particle.

The brass stack proves to be a very efficient detector for soft showers of energy above 10^{11} ev and for the primaries initiating these events. It should therefore be possible to extend the energy spectrum for protons and α-particles to energies above 10^{13} ev.

We wish to express our gratitude to Bausch and Lomb Optical Company, who constructed for us a special microscope which was indispensable for some of the measurements described here.

We are also indebted to the staff of General Mills Aeronautical Laboratory and especially to Mr. A. T. Bauman, who made the balloon flights.

The largest and hardest part of the scanning work was done by Mrs. Katherine Reynolds, whose contribution was invaluable. We are also indebted to Mr. N. Wilson, Mrs. E. Woodruff, and Miss H. Pedell, who assisted in the scanning.

We wish to thank the Office of Naval Research for arranging the balloon flights.

APPENDIX I. STATISTICAL TREATMENT OF CENTER OF GRAVITY AND PLATE SPACING CORRECTIONS

In performing measurements on showers it was usually not possible to determine experimentally the axis of the shower. In these cases the "center of gravity" of the particles was used as an axis. Section A relates the opening angle measured relative to the center of gravity to the true opening angle. Section B relates the scattering measured relative to the center of gravity to the true scattering. Section C concerns the more general problem of removing simulated scattering caused by relative rotation, tilting and uneven spacing of the plates. We derive the relation which was used to correct for uneven plate spacing since this was the only cause of error large enough to affect the measurements in the stack used. The identical method can be used to derive the more general relation. At a distance R from the origin of the shower we construct a plane perpendicular to the shower axis. We choose as the origin of our coordinate system in this plane its intersection with the shower axis.

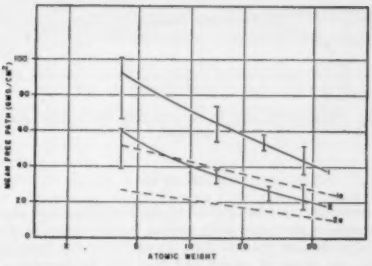


FIG. 13. Mean free paths in brass (1) and glass (2) for particles of various atomic weights. The solid curves represent the empirical relation of Eq. (11), ($\Delta R = 1.0 \times 10^{-18}$ cm) and the dotted curves are the corresponding geometric mean free paths ($\Delta R = 0$).

We define the following symbols:

$X_{i,\lambda}(\lambda = 1,2)$ are the Cartesian coordinates of the intersection of the ith track with the plane.

$\theta_i = (1/R)(\Sigma_\lambda X_{i,\lambda}^2)^{\frac{1}{2}}$ is the angle of the ith track with the shower axis.

$X_\lambda = (1/n)\Sigma_i X_{i,\lambda}$ are the coordinates of the intersection of the center of gravity axis with the plane.

$x_{i,\lambda} = X_{i,\lambda} - X_\lambda$ are the coordinates of the intersection in the plane of the ith track measured with respect to the intersection of the center of gravity axis.

$\phi_i = (1/R)(\Sigma_i x_{i,\lambda}^2)^{\frac{1}{2}}$ is the angle between the ith track and the center of gravity axis.

$S_{i,\lambda}^T$ is the true scattering of the ith track projected on the λ-axis.

$S_{i,\lambda}$ is the measured scattering or the second differences in the position of the ith track with respect to the center of gravity of the shower.

n is the total number of shower particles.

Section A

The mean square opening angle of the shower tracks with respect to the shower axis ($\langle\theta^2\rangle$), can be expressed in terms of the measurable mean square opening angle of the tracks with respect to the center of gravity of the shower ($\langle\phi^2\rangle$).

In terms of the symbols defined above we have:

$$x_{i,\lambda}^2 = (X_{i,\lambda} - X_\lambda)^2 = [X_{i,\lambda} - (1/n)\sum_j X_{j,\lambda}]^2,$$

$$\sum_i x_{i,\lambda}^2 = \sum_i X_{i,\lambda}^2 - \frac{2}{n}\sum_i X_{i,\lambda}^2 + \frac{1}{n}\sum_i X_{i,\lambda}^2 = \frac{n-1}{n}\sum_i X_{i,\lambda}^2,$$

since for tracks randomly distributed in azimuth

$$\sum_{i,j} X_{i,\lambda} X_{j,\lambda} = 0 \text{ for } i \neq j.$$

Using the definitions given above we therefore obtain

$$(1/n)\sum_i \phi_i^2 = [(n-1)/n^2]\sum_i \theta_i^2 \quad \text{or} \quad \langle\phi^2\rangle = [(n-1)/n]\langle\theta^2\rangle.$$

FIG. 14. Parameters used in determining the sagitta scattering constant.

Section B

A correction must be applied to the scattering data since the measurements are made with respect to the center of gravity. Assume for the moment that there is a rigid reference line travelling through the stack of plates and that the displacements with respect to it are $X_{i,\lambda}$ for the first plate, $X_{i,\lambda}'$ for the second plate, and $X_{i,\lambda}''$ for the third plate. The second difference of the displacements from the center of gravity is

$$S_{i,\lambda}=(X_{i,\lambda}-2X_{i,\lambda}'+X_{i,\lambda}'')+(2\sum_j X_{j,\lambda}'-\sum_j X_{j,\lambda}-\sum_j X_{j,\lambda}'')/n.$$

The true scattering is

$$S_{i,\lambda}{}^T=(X_{i,\lambda}-2X_{i,\lambda}'+X_{i,\lambda}'').$$

Therefore

$$S_{i,\lambda}=S_{i,\lambda}{}^T-(1/n)\sum_j S_{j,\lambda}{}^T.$$

Squaring and taking the average (the procedure is identical to that of A)

$$\langle S^2\rangle=[(n-1)/n]\langle(S^T)^2\rangle,$$

where the average of the cross terms is zero since there is equal chance for $S_{i,\lambda}{}^T$ to be positive or negative.

Section C

A statistical correction which removes the simulated scattering due to plate spacing differences is possible because the plate spacing differences change the distance between tracks by the same percentage. The process consists of minimizing the measured scattering by hypothetically changing the plate separation and then adding back the amount of true scattering removed in this process. The derivation is as follows: The measured scattering is

$$S_{i,\lambda}=S_{i,\lambda}{}^T-(1/n)\sum_j S_{j,\lambda}{}^T-ax_{i,\lambda},$$

where a is a correction term which compensates for the difference in plate spacing since the simulated scattering due to plate spacing differences is proportional to the projected opening angles of the tracks.

We define:

$$\Omega_{i,\lambda}=S_{i,\lambda}-a'x_{i,\lambda},$$

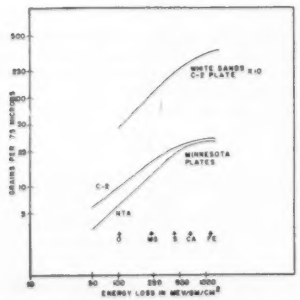


FIG. 15. Calibration curves for underdeveloped C-2 and NTA emulsions.

where a' has such a value that $\sum_{i,\lambda}\Omega_{i,\lambda}{}^2$ is a minimum. The value of a' is then

$$a'=\sum_{i,\lambda}x_{i,\lambda}S_{i,\lambda}/\sum_{i,\lambda}x_{i,\lambda}{}^2.$$

With the above definitions

$$\Omega_{i,\lambda}=S_{i,\lambda}{}^T-\frac{1}{n}\sum_j S_{j,\lambda}{}^T-x_{i,\lambda}\sum_{j,\mu}x_{j,\mu}S_{j,\mu}{}^T/\sum_{j,\mu}x_{j,\mu}{}^2.$$

Since $\sum_j x_{j,\lambda}=0$ by definition, we find on squaring and taking the average that

$$\sum_{i,\lambda}\Omega_{i,\lambda}{}^2=(2n-3)\langle(S^T)^2\rangle$$

or

$$\langle\Omega^2\rangle=[(2n-3)/2n]\langle(S^T)^2\rangle.$$

APPENDIX II. SCATTERING CONSTANT

In order to determine the scattering constant we note that (a) the measurement is a second difference measurement (sagitta type), and (b) the scattering takes place across cells composed of layers of glass, brass, emulsion, and air.

It can be shown that the sagitta scattering constant is simply related in the Gaussian approximation to the Williams-Molière theory of scattering between successive tangents in a homogeneous medium. In Fig. 14 the second difference is $S=Y_1+Y_2$, where Y_1 and Y_2 are the distances of the track from the tangent to the track at the center of the cell and $\langle S^2\rangle=\langle Y_1{}^2\rangle+\langle Y_2{}^2\rangle$. The formula for Y^2 is given by $t^3\alpha_3{}^2/3$, where $\alpha_3{}^2$ is the Williams scattering parameter for a unit cell length. In an inhomogeneous medium

$$\langle Y^2\rangle=\tfrac{1}{3}t^3\{\alpha_1{}^2[1-(1-k_1)^3]+\alpha_2{}^2[(1-k_1)^3-(1-k_1-k_2)^3]+\cdots$$
$$+\cdots\alpha_m{}^2(1-\sum_{i=1}^{m-1}k_i)^3\},$$

where $\alpha_m{}^2$ is the scattering parameter for the various materials in order of succession from the center of the cell and k_i is the fraction of the cell taken up by each material. Using these expressions for the brass stack the scattering relations for particles with $Z/m=\frac{1}{2}$ become

$$\langle S^2\rangle^{\frac{1}{2}}=14.5t^{\frac{3}{2}}/p\beta \text{ for a 1 plate spacing,}$$
$$\langle S^2\rangle^{\frac{1}{2}}=15.4t^{\frac{3}{2}}/p\beta \text{ for a 2 plate spacing,}$$

where S is in microns, t is in units of 5000 microns, $\beta=v/c$, and p is the momentum in units of the proton rest mass. For the particular cell lengths chosen, the values of α^2 have the ratios

$$\text{brass:emulsion:glass}=9.7:4.6:1.$$

APPENDIX III. CHARGE CALIBRATION

At northern latitudes a large fraction of the particles are slow (nonrelativistic). Therefore track identification must be based on simultaneous ionization and range measurements.

The NTB-3 and G-5 emulsions were developed with the usual temperature method. The C-2 and NTA emulsions were heavily underdeveloped using the usual temperature method with D-25 as the developer. The warm stage of the process occurred at 15°C for 15 minutes.

The calibration of the insensitive plates was established in the following manner: (a) In the insensitive plates, tracks of α-particles emitted from stars were observed and grain counted at known distances from the end of their range. To assure that these particles were alphas, the end of their ranges were compared with the alphas from thorium stars. This calibrated a range of ionization from 60 to 500 times minimum. (b) The heaviest ionizing tracks occurring with reasonable frequency and penetrating the stack are due to iron nuclei. This fact was used to establish the high ionization end of the calibration curve for the insensitive plates.

Thus a complete calibration for the insensitive plates was obtained from an ionization corresponding to relativistic iron to one

corresponding to relativistic oxygen. The δ-ray calibration was established as follows:

(a) High energy α-particles (approximately 20 Bev per nucleon) resulting from a *C*-collision, were δ-ray counted. Minimum ionization tracks were also δ-ray counted. Since for relativistic particles:

$$\delta_E = Z^2 a + b,$$

where "*a*" is a constant and "*b*" the δ-ray density due to random electrons, both "*a*" and "*b*" can be determined.

(b) Because of the difficulty in maintaining the same criterion for δ-ray counting when counting a track of minimum ionization the method described above is not extremely reliable. Therefore particles which were identified as oxygen in the insensitive plates were δ-ray counted in the sensitive plates. The results were consistent with the δ-ray calibration from relativistic protons and α-particles.

Using this calibration we obtain in the primary flux a strong peak corresponding to the CNO group and very few particles of charge $2<Z<6$ in agreement with our previous results.[3]

The calibration curve for the C-2 and NTA emulsions is shown in Fig. 15. It will be noted that the curves reach saturation at very low grain densities. We presume this could be avoided by using less sensitive emulsions and less underdevelopment. For steep heavy tracks δ-ray counts became unreliable and grain counts were used to determine charge.

Corresponding results were obtained with the White Sands charge calibration.

PHYSICAL REVIEW VOLUME 85, NUMBER 2 JANUARY 15, 1952

The Variation of Intensity of Fast Cosmic-Ray Neutrons with Altitude

L. F. CURTISS*
National Bureau of Standards, Washington, D. C.

AND

P. S. GILL
Department of Physics, Aligarh University, Aligarh, India

(Received August 29, 1951)

Using BF_3 proportional counters, containing boron enriched in B^{10}, imbedded in a paraffin block 15×15 ×17 inches, as a moderator of fast neutrons, the intensity of such neutrons generated in the atmosphere by cosmic rays has been measured at a number of altitudes between 5000 and 13,000 feet above sea level. These locations were all along the same geomagnetic latitude of 20° 32′ N in Kashmir. The intensity thus measured was found to increase approximately in an exponential manner with decreasing atmospheric pressure under the conditions of the experiments, in general agreement with observations made by other investigators in the free atmosphere, giving an attenuation length in the atmosphere of 128 g/cm².

I. INTRODUCTION

CONSIDERABLE interest has been shown in the last decade in the neutron component of cosmic radiation with attention focused chiefly on the slow neutrons in the atmosphere. The slow neutrons offer the advantage of detection by relatively simple means, such as BF_3 counter tubes, which makes possible the use of equipment of light weight suitable even for free balloon flights. This accounts in part for the numerous investigations of their variation in intensity with altitude.

Detectors of fast neutrons, on the other hand, are usually heavy and bulky, since to date the most satisfactory method of detecting these neutrons is to moderate them in some medium, such as paraffin, and then to detect them as slow neutrons with BF_3 counters imbedded in the moderator. Such devices cannot be carried aloft readily by balloons. Consequently about the only convenient way of making measurements of fast neutrons generated in the atmosphere at various levels is either to transport the equipment to various heights on the surface of the earth or to use airplanes. Observations with airplanes suffer from an inability to maintain even an approximately constant position laterally and vertically for the relatively long time required to obtain measurements of a statistical accuracy of the order of one percent. Also the use of planes for this purpose is expensive. Since fast neutrons are presumably a direct product of that type of interaction between the cosmic radiation and atomic nuclei which leads to nuclear disintegration, it is reasonable to suppose that the study of fast neutrons in the atmosphere would give more direct information regarding the nature of this process than is obtained from data regarding slow neutrons.

These considerations led the authors to attempt to obtain information regarding the variation of intensity of fast neutrons in the atmosphere over a limited range of altitudes at some elevated location on the surface of the earth. The region selected for these trials is in Kashmir where altitudes between 5000 and 13,000 feet are accessible lying along the same geomagnetic latitude 20° 32′ N. Also adequate manpower at low cost is available there to carry the relatively heavy equipment up over the snow fields and rough terrain in the steep

* Fulbright research scholar assigned to Aligarh University, Aligarh, India.

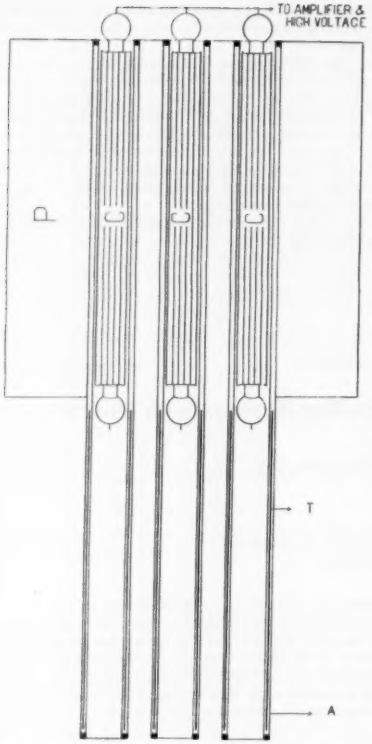


FIG. 1. Plan of detector block; *P*—Paraffin block; *C, C, C*—BF_3 proportional counters; *T*—Tubular cadmium shields; *A*—Aluminum guide tubes which support cadmium shields when withdrawn.

ascents encountered. This paper describes the equipment, methods of observation, and results obtained on such an expedition in Kashmir from May to August, 1951. Two separate ascents were made in the Himalayan Mountains from Srinagar, the capital of Kashmir. One was through Tangmarg, Gulmarg, and Khellanmarg to Alapathar Lake. The other went through Pahalgam, Goshangan, and Lower Katar Nag to Katar Nag near the Kolahoi Glacier. The maximum altitude attained in each case was about 13,000 feet.

II. EXPERIMENTAL EQUIPMENT

Although it has been usual to employ battery-operated equipment in making cosmic-ray studies in the field, it was decided in this case to use a more or less conventional laboratory amplifier and scaler designed to operate from a 110-volt ac supply. This procedure permitted the use of equipment which had been tested over long periods in the laboratory and could therefore be expected to be thoroughly reliable. On the expedition, this apparatus was operated from a gasoline-driven generator giving 110 volts at 60 cycles and capable of delivering 500 watts. The outfit selected had a centrifugal governor on the engine, permitting the voltage to be adjusted quite accurately and also holding this voltage at a remarkably constant value under the constant load represented by the electronic equipment.

Three BF_3 proportional counters were used to detect the moderated neutrons. These counter tubes were made of brass with Kovar seals in each end and were 14 inches long and 1.5 inches in external diameter. They were of commercial manufacture, and the boron used in the filling mixture was about 98 percent B^{10}. The pressure of the filling gas was about 45 cm and the operating voltage extended from 2700 to 2900 volts with less than ± 1 percent change in counting rate. This range was nearly identical for each tube. These counter tubes were mounted in a paraffin block 15 by 15 by 17 inches, with centers separated by 3 inches as shown in Fig. 1. They were supported inside aluminum tubes passing through the paraffin block and extending to one side a distance slightly greater than the length of the counter tubes. These aluminum tubes served as guides for tubular cadmium shields which could be slid over the counter tubes to exclude thermal neutrons and thus measure the background of the counters. This procedure was also useful in checking the proper operation of the equipment. Any defects which might cause spurious counts could be readily detected, if present, with these shields covering the counter tubes. In this way leakage of high voltage over insulation or corona discharge would have been revealed. Little trouble of this kind was encountered. These cadmium shields cut out all but about 2 percent of the slow neutrons reaching the counter tubes. This was determined by measurements with a Ra-Be α-type source, containing about 100 mg of radium. Thus the shields reduced the normal counting rate to a correspondingly low value. The actual background of the counter tubes, which should have the same average value throughout the experiments, was taken as the average counting rate at sea-level with the tubes covered with cadmium. This was found to be 0.15 ± 0.02 count per minute under the conditions maintained throughout the observations. The measured background count increased with altitude, roughly proportional to the increase in total neutron count. This slight increase was attributed to the previously mentioned leakage of neutrons past the cadmium shields, rather than to a direct effect on the counter tubes of some component of the cosmic radiation. The maximum background rate at the higher altitudes was approximately 0.65 count per minute. It is to be noted that this detector for fast neutrons responded to those generated in the paraffin block and in the earth nearby, as well as in the surrounding atmosphere.

The high voltage for operating the proportional counter tubes was obtained from the rectifier circuit in the commercial scaler used for recording the counts. This circuit delivered a maximum of 2500 volts dc. Since the tubes required 2800 volts, a 300-volt miniature battery was inserted in the high voltage line to the counter tubes. The pulses from these three tubes, connected in parallel, were amplified by a Jordan and Bell amplifier,[1] and fed to the scaler. This amplifier contained a discriminator circuit which was adjusted in combination with the gain control to give most satisfactory operation. Once these conditions were ascertained, by tests in the laboratory with a neutron source, neither control was changed throughout the experiments. It was found desirable to reject about 10 percent of the actual neutron-induced pulses in order to make sure that spurious pulses were not recorded by the scaler. This small percentage of the neutron pulses was found to have heights only slightly greater than the noise level of the amplifier, so that it was impractical to attempt to record all pulses. Since the same fraction was rejected for all observations by a pulse-height discriminator attached to an appropriate stage of the amplifier, this procedure could not be expected to interfere seriously with the measured relative values of neutron intensity. Each component was crated in a wooden box, with adequate padding, to be carried by coolies from one experimental site to another.

III. METHOD OF OBSERVATION

It was decided at the start of the series of measurements to arrange the components in a definite way which was maintained throughout the observations. The detector block was placed directly on the ground, the amplifier housed in a heavy steel cabinet on top of this block with the scaler resting on the amplifier cabinet. This arrangement required a minimum of ground space in the tent and was convenient for operation. To protect the apparatus from direct sunlight and from rain, it was sheltered in a small tent which was always placed at some little distance from buildings or other obstructions, such as banks of snow or walls of rock. Tests made during the expedition revealed that it was important to adhere rigidly to the definite arrangement of parts and to avoid proximity to obstructions. Any changes in arrangement or placing the equipment near obstructions caused changes in the apparent efficiency of the detector.

When in operation, the gasoline-driven generator was monitored for output voltage at frequent intervals. However, the governor on the engine was so efficient that this was in general only a precautionary measure. Drifts of voltage after the initial warm-up period were rare and momentary fluctuations of voltage did not exceed ± 1 percent in normal operation.

Readings of the scaler and of a large calibrated aneroid barometer, compensated for the effect of temperature changes, were made every hour during observation periods. The barometric pressure assigned to a given point was the average of these hourly readings over the time required to accumulate at least 10,000 neutron counts. Usually there was little change in the barometric readings for a given set of observations. Measurements with the cadmium shields over the counter tubes were made at intervals to check the operation of the circuits. Since the counting rates varied from about 4 per minute at 5000 feet to roughly 18 per minute at 13,000 feet, the time required for a complete set of observations ranged from about 60 to 15 hours.

While at Gulmarg, advantage was taken of the opportunity to observe the effect of an additional slab of paraffin, 4 inches thick and of an area equal to that of one side of the detector block, when this slab was applied to different sides of the block. The effect in all cases was to reduce the counting rate by about 10 percent, with little if any indication of any directional effects in the neutron flux. Furthermore, when the equipment was brought inside a frame house, a similar reduction in the apparent neutron intensity was noted. This confirms the suggestion, made by Cocconi,[2] that the neutrons generated in the atmosphere have energies of less than 10 Mev. Otherwise this additional paraffin might have been expected to increase the efficiency of the detector block.

IV. EFFECT OF TEMPERATURE CHANGES

Temperatures of the equipment changed quite radically during a 24-hour period of observation. When the sun was shining the temperature of the interior of the tent rose to about 90°F in the early afternoon, and it fell to about 35° to 40°F at most sites in the early hours of morning. It was impracticable to measure the temperatures of the detector block under field conditions, and it was hoped that the effect of changes in temperature, if these were marked, would average out over the relatively long observation period at each site. It was further suspected that any effect of temperature change was quite small. This was deduced from comparing hourly counting rates taken at various times during the day which were found to be quite constant, within limits of statistical accuracy.

To answer more completely the question of the effect of temperature change, a metal box was constructed on returning to Aligarh University in which the detector block was placed. A surrounding wooden box was used to retain a layer of ice in contact with the metal box, and counting rates with a Ra-Be neutron source were observed at different temperatures measured by a thermometer imbedded in the paraffin of the detector block. This study of the effect of change of temperature was complicated by the fact that only a relatively

[1] W. H. Jordan and P. R. Bell, Rev. Sci. Instr. **18**, 703 (1948).

[2] V. T. Cocconi, Phys. Rev. **74**, 226 (1948).

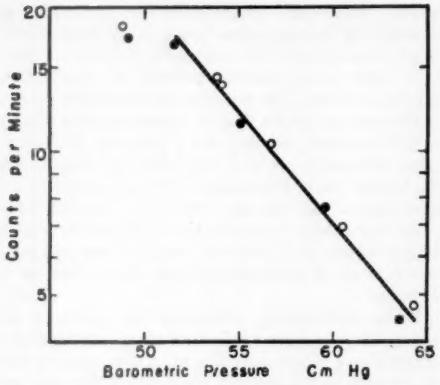


FIG. 2. Intensity of fast cosmic-ray neutrons vs barometric pressure. Open circles taken on Srinagar, Tangmarg, Gulmarg, Khellanmarg, and Alanathar Lake ascent. Solid circles taken on Srinagar, Pahalgam, Goshangan, Lower Katar Nag, and Katar Nag ascent.

strong source of neutrons (100 mg Ra) was available for the test, requiring the source to be outside the cooling arrangement. Under these conditions the neutron counting rate was dependent on the density of the layer of ice used for cooling. When precautions were taken to keep the same amount of ice in this layer, it was found that in the range of 13.1°C to 35.9°C the counting rate was not noticeably affected by differences of temperature. In view of the fact that no marked effect of temperature would be predicted for this detector block, this was regarded as satisfactory proof of the absence of such effects within the one percent accuracy aimed for in the measurement of the fast cosmic-ray neutrons.

V. EXPERIMENTAL RESULTS

The counting rates, corrected for the small background rate, obtained at nine different sites in the two ascents are plotted semilogarithmically in Fig. 2, against the corresponding average barometric pressure. The observations, with the exception of one point for each ascent, follow an exponential curve with a surprisingly small scattering of points considering the nature of the measurements and the conditions under which they were taken. As has been indicated, the two ascents were in localities about 80 miles apart and about a month intervened between these series of measurements. The peculiar fall of intensity at the uppermost point is evident in each series and therefore might be regarded as real. This deviation is difficult to explain on the basis of any defect in the equipment since leakage of voltage over insulators or corona discharge in the reduced air pressure would almost certainly be revealed by a marked increase in the counting rate. Apart from these two points, the points lie closely along a straight line with a slope corresponding to an attenuation length in the atmosphere of 128 g/cm² which is approximately that quoted by Rossi[3] for cosmic-ray bursts and close to that for stars and penetrating showers. It is considerably less than the value for slow neutrons. This is to be expected since the processes recorded as bursts and stars are assumed to give rise to the fast neutrons in the atmosphere. In general it can be concluded that measurements made at various altitudes on the earth, as described in this paper, give results quite similar to those obtained in the free atmosphere with balloons or planes under more convenient conditions of observation.

If measurements of the rate of production of fast neutrons in the atmosphere prove to have value in exploring the nuclear reactions which accompany cosmic radiation, and in particular the study of the variation of this rate with altitude, the present experiments may serve as a preliminary guide in planning similar observations in the future, where the experiments are to be performed on the surface of the earth. They may also be helpful in calling attention to some of the precautions to be taken in such studies and serve to emphasize the necessity for observations of relatively high statistical accuracy. Obviously some of the scattering of the points in Fig. 2 must be attributed to the actual fluctuation of intensity of the cosmic radiation itself. It is not easy to see how such effects can be avoided or their influence on the data compensated except by extending observations over much longer periods than was possible in the present experiments.

VI. ACKNOWLEDGMENT

In conclusion the authors wish to acknowledge the support of the Fulbright Program which made it possible for one of us (L.F.C.) to come to India, and to thank Dr. E. U. Condon, Director of the National Bureau of Standards, for permission to participate in this work. We are also indebted to Aligarh University for financial aid and appreciate in particular the interest which the Vice-Chancellor, Dr. Zakir Husain, has taken in this work. This work was greatly facilitated by the cooperation of various officials of the Kashmir Government and by the aid of several graduate students and staff members of the Department of Physics at Aligarh University, who gave time from their summer vacation to help in taking the observations under what were often difficult conditions.

[3] B. Rossi, Revs. Modern Phys. 20, 537 (1948).

The Absorption of Selenium in the Soft X-Ray Region*

M. Parker Givens and Walter P. Siegmund
Institute of Optics, University of Rochester, Rochester, New York
(Received September 25, 1951)

The absorption of selenium (from 180A to 240A) has been investigated experimentally. The absorption edge near 225A has been resolved into a double edge due to M_{IV} and M_V absorption. The separation of these edges is 0.7 ev. A minimum in the absorption is observed at 208A.

I. INTRODUCTION

In the process of x-ray absorption by solids the energy of the light quantum $h\nu$ is used to raise an electron from its initial energy state into another state which is both permitted and empty. The probability that an electron will absorb the energy of an incoming quantum and be raised to a vacant state of energy E is proportional to the density of permitted states with energy near E. This quantity is usually called $N(E)$. The ultimate goal of most absorption measurements is to furnish information which may aid in the eventual determination of $N(E)$ for the substance under study.

The soft x-ray region (or the extreme vacuum ultraviolet) is suitable for this work because the electrons involved are lifted into the outer bands from initial levels which are sufficiently narrow that their width may be ignored. At the same time the total energy involved in the transition is considerably less than 100 times the width of the upper energy band. For this reason the spectrometer need not have unusually high resolving power.

II. EXPERIMENTAL PROCEDURE

Figure 1 is a simplified diagram of the equipment. The Rowland circle is one meter in radius. The plate holder follows the circle and the circle passes through the slit. The grating of 30,000 lines per inch is tangent to the Rowland circle. The angle of incidence is 85°. The source used for this work was a vacuum spark between copper electrodes operated from a 50-kv power supply. The equipment is essentially the same as that used by Skinner and Johnston[1] and Sanner.[2] Wavelengths were determined from known lines[3] in the copper spectrum.

The absorber was mounted on a small metal plate with a hole in it, and this plate was located just behind the slit.

The absorber consisted of a thin foil of selenium evaporated onto a thin film of celluloid. In order to obtain the absorption due to selenium it is sufficient to compare the transmission through two absorbers for which the thickness of the selenium is different, but which are otherwise identical. Neither is it wise to use a zero thickness of selenium on one of the absorbers, since this would alter the reflection of the radiation from the boundaries. Also, Pell and Tomboulian[4] have observed that for aluminum there is an extra absorption band associated with the interface between the metal and the backing material. The use of a comparison absorber with no selenium would result in attributing such a band (if it exists for a collodion-selenium interface) to the selenium. On the other hand, if both absorbers have a sufficient thickness of selenium any effect associated with the interface will not show up in the comparison. For these reasons no measurements were made from absorbers having less than 470A of selenium.

The two absorbers used in any given comparison were made with celluloid from the same floating so that the celluloid films would be as identical as possible. These two were placed in the evaporator together and selenium deposited on one to a thickness t_1, with the other covered over by means of a plate which could be moved from outside the vacuum. The plate was then turned aside and the evaporation continued until the selenium on the second absorber reached a thickness t_2. The thicknesses t_1 and t_2 were determined by the interference of visible light reflected from the two surfaces of the selenium. The conversion to actual thickness requires the use of the index of refraction of selenium for the visible light used. Mercury green light was used and the index was found to be 2.9. The celluloid backing films were approximately 300A thick.

In a normal run four exposures were made on the

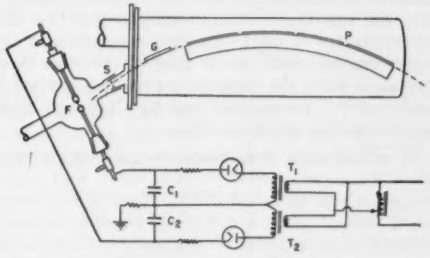


FIG. 1. The vacuum spectrograph showing the plateholder (P), the grating (G), the slit (S), and the vacuum spark (F). The essential features of the power supply are also shown.

* This work was supported by the ONR.
[1] H. W. B. Skinner and J. E. Johnston, Proc. Roy. Soc. (London) A161, 420 (1937).
[2] H. Sanner, Z. Physik 94, 523 (1935).
[3] P. G. Kruger and F. S. Cooper, Phys. Rev. 44, 826 (1933).
[4] D. H. Tomboulian and E. M. Pell, Phys. Rev. 83, 1196 (1951).

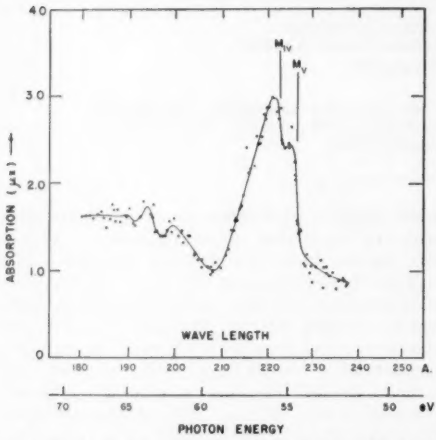


FIG. 2. The absorption due 1900A of selenium. Notice the double edge near 225A and the spectral window at 208A.

same photographic plate. Three of these were made through the thinner absorber with varying exposures of K_1, K_2, and K_3. The fourth exposure was then made through the thicker absorber with an exposure of K_4. The exposure was determined by counting the number of sparks produced by the source. The plate was then developed with good agitation so that all four exposures received identical processing. Densitometer traces were then made from these spectra and measured in the usual manner.

The first three exposures are used to determine the response curve of the photographic emulsion. At each wavelength we have three points on the curve except that the exposures are relative not absolute. Using a number of different wavelengths many sets of points can be obtained, and these are then fitted into a smooth response curve by matching exposures at overlapping densities. The response curve obtained in this way is based on the assumptions that the response is constant over the wavelength range employed and that the reciprocity law is valid. The final comparison of the fourth exposure with one or more of the other three was made using this response curve. This method is reminiscent of the method used by Cady and Tomboulian[5] in their emission studies.

[5] W. M. Cady and D. H. Tomboulian, Phys. Rev. **59**, 381 (1941).

III. RESULTS AND CONCLUSIONS

The results of these investigations are shown in Fig. 2 which shows the absorption of selenium against wavelength and photon energy. The data for this curve were taken from two absorbers for which t_1 (the difference in thickness of selenium) was 1900A and t_2 was 480A. The measurements were repeated using another (but similar) pair of absorbers. The new curve and the original were practically identical in shape and location in the spectrum, whereas the values of μx agreed within about 5 percent. For the curve shown $K_1=4000$; $K_2=2000$; $K_3=1000$; and $K_4=4000$.

The absorption edge at 225A is seen to be double representing the separation of the M_{IV} and M_V edges. This separation of 3A or 0.7 ev is consistent with the value obtained by extrapolating from the M_{IV-V} separation of tungsten using the rule that this separation is proportional to $(Z-13)^4$.

The magnitudes of these two absorption jumps are predicted to be in the ratio of 2:3. The ratio is observed to be about 1:4, the M_V jump being relatively larger than predicted. Skinner has observed that in emission the L_{III}/L_{II} intensity ratio is the predicted value of 2 only for insulators. For semiconductors and metals the ratio is much larger. He attributes this to an Auger process. It is difficult to visualize how such a process could alter the strength of the absorption edges.

There is a minimum in the absorption about 4.6 ev above the absorption edge. Since this is essentially the same as 4.62 ev reported[6] as the work function of selenium, this minimum corresponds to electrons which are nearly freed from the solid by the photon. Since the selection rule $\Delta L=\pm 1$ indicates that transitions $d\rightarrow s$ and $d\rightarrow d$ should be improbable, the absorption curve for the M_{IV-V} electrons represents the density of states $N_p(E)+N_f(E)$. Also, $N_f(E)$ for selenium is small so the absorption largely represents $N_p(E)$. The decline in absorption for wavelengths shorter than 218A is probably due to a decrease in the density of these states. The fact that the conduction band arises from p electrons, whereas the next two bands above it arise from s and d electrons, gives support to this viewpoint.

The structure at shorter wavelengths is without explanation since the energies are too small to give Kronig structure, and the next absorption edge (M_{II-III}) is expected to lie near 100A.

[6] R. Hamer, J. Opt. Soc. Am. **9**, 256 (1924).

Ferroelectricity and Antiferroelectricity of a Crystal Containing Rotatable Polar Molecules*

YUTAKA TAKAGI
Tokyo Institute of Technology, Oh-okayama, Tokyo, Japan
(Received August 24, 1951)

We have investigated the dielectric properties of a CsCl type crystal made up by insertion of ions into the body center sites of a simple cubic array of rotatable polar molecules. Owing to the Lorentz field exerted by the dipoles themselves, a non-uniform distribution in orientation of the polar molecules is predicted below a certain critical temperature. The main results obtained are as follows: (1) It is shown that the crystal is antiferroelectric so long as the polarizability of the central ions is small, half of the dipoles being partially oriented parallel to and the other half antiparallel to one of the cube edges. On the other hand, if the polarizability of the ions is larger than a certain threshold value, the crystal is ferroelectric, all the dipoles being partially oriented parallel to each other. (2) The dielectric constant-temperature curves calculated for antiferroelectric cases show a variety of forms according to the polarizabilities of both components. For some specimens they are almost flat throughout the region around their own critical temperatures, while for other specimens they can show such sharp peaks that they are hardly discernible from those of ferroelectric crystals, which always show very sharp peaks tending to infinity. (3) Below their respective critical temperatures, the dielectric constant should decrease with increasing dc bias in a ferroelectric crystal, while in an antiferroelectric crystal it should increase slightly under the same condition. (4) If a sufficiently strong field is applied to an antiferroelectric crystal at just below its critical temperature, it will be forced to make a momentary transition to a ferroelectric one, the change being accompanied by an abrupt increase of polarization.

It is suggested that most of the phenomenological predictions derived with this model may, qualitatively at least, be valid also for other models which have ions capable of displacement instead of freely rotatable dipoles, as is really the case with the well-known $BaTiO_3$ and the like.

I. INTRODUCTION

ANOMALOUS peaks are frequently observed in the dielectric constant-temperature curves of crystals containing or composed of rotatable polar molecules, and are generally known as the phase changes arising from hindered molecular rotation.[1] Considering the form of the lambda-points of the peaks, the transitions must represent some kind of cooperational phenomenon, and the hindering potential acting upon any molecule on the lattice must in any case depend upon the state of rotation of the molecules themselves.

Several attempts have been made to construct an explicit form for the torque potential. For example, Kirkwood[2] has formulated a theory based upon a short-range interaction of the form $(\epsilon/2)\cos\gamma$ for a pair of neighboring molecules, the axes of which are inclined at an angle γ to each other. If the hindering potential in the crystal concerned really arises from short-range forces, such as steric hindrance or the like, which can be effective only between the molecules in contact, his treatment will yield a good approximation. On the contrary, however, if it is certain that most of the hindering torque arises from the dipole-dipole interaction, as it does in the case with which we are going to deal, the nearest-neighbor treatment may be inadequate, as Kirkwood himself mentioned in his article. This is, in the first place, because the dipole forces are long range in nature and, what is still more important, because their peculiar directional nature invalidates the expression in terms of a single parameter γ. For example, two dipoles parallel to each other have a negative energy if they are on the same line, whereas they can have even a positive energy if they are arranged side by side.

The present paper aims to show how a fairly satisfactory theory can be formulated if we replace Kirkwood's potential by the Lorentz field. For the sake of simplicity, we shall start with a simple cubic array of rotatable dipoles, for we know well what is the most suitable set of sublattices into which a simple cubic lattice can be decomposed.[3] Besides, we shall insert ions into the body center sites of the array in order to clarify an important role played by the polarizabilities of the crystal components. It will be shown how interesting characteristics unknown in magnetic substances are theoretically expected for dielectric crystals only.

Very recently Kittel[4] has developed a theory of antiferroelectricity. The present theory will add to it from another point of view, though the model treated here is perhaps too simplified to be realized in nature.

II. DISTRIBUTION FUNCTIONS

To start with, let us consider a simple cubic array of dipoles without the ions at the body center sites. Electrostatic energy eigenvalues of such an array have been fully discussed by Luttinger and Tisza.[3] They have shown that in the state of lowest energy which the system should assume at 0°K, half of the dipoles, namely those on the lattice denoted by 1 in Fig. 1, orient themselves parallel to, say, the $+z$ direction while the other half, on the lattice denoted by 2, orient themselves parallel to the $-z$ direction.

* The expense of this research has been defrayed from the Scientific Research Expenditure of Japanese Ministry of Education. A preliminary account of the present paper was given at the 1950 Annual Meeting of Phys. Soc. Japan, held at Osaka University, on November 4, 1950.

[1] R. H. Fowler, *Statistical Mechanics* (Cambridge University Press, London, 1936), second edition, p. 810.

[2] J. G. Kirkwood, J. Chem. Phys. 8, 205 (1940).

[3] J. M. Luttinger and L. Tisza, Phys. Rev. 70, 954 (1946).

[4] C. Kittel, Phys. Rev. 82, 729 (1951).

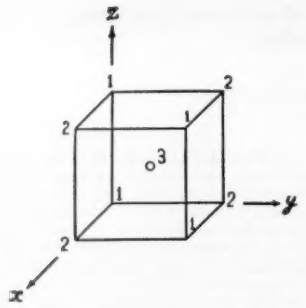


FIG. 1. Classification of sublattices. The sites 1 and 2 are occupied by polar molecules; the site 3 by ions.

For such an ordered configuration, full use of symmetry considerations may be made in calculating the inner field. It turns out that the field acting on lattice 1 is really pointing in the $+z$ direction while that on the lattice 2 points in the $-z$ direction.

As the temperature is raised, however, this ordered configuration will be gradually disturbed by thermal agitation, and a rigorous calculation of the inner field becomes very difficult. That is, although we are sure that the inner field will then decay in some kind of cooperative manner, yet a rigorous evaluation of it seems almost impracticable because one must carry out a summation over a great many electrostatic forces, each arising from an individual dipole oriented at random.

Viewed from another angle, however, this very process of summation seems at the same time to have the advantage of smoothing out most of the microscopic fluctuations in the local field from site to site. So we may assume without much harm that the inner field F_i acting on a lattice i is homogeneous throughout the specified lattice, though of course it can be different from the inner field F_j acting on another lattice j; its value, to this approximation, may be calculated by supposing that every molecule had just the average moment taken over the specified lattice to which it belongs. That is, we shall assume that the field is to be determined if we only know the macroscopic polarizations P_i of the sublattices.[5,6]

Now we shall add the ions into the body center sites and designate them by the number 3. On account of our assumption

$$F_i=E+\sum_{j=1}^{3} f_{ij}P_j/\epsilon_0; \quad i=1,2,3, \tag{1}$$

where f_{ij} are the Lorentz factors, the values of which may be obtained from a paper by McKeehan;[7]

$$\begin{aligned} &f_{ij}=f_{ji}, \\ &f_{11}=f_{22}=\tfrac{1}{3}+0.4260, \\ &\quad f_{12}=\tfrac{1}{3}-0.4260, \\ &f_{13}=f_{23}=f_{32}=\tfrac{1}{3}. \end{aligned} \tag{2}$$

The polarizations P_i are in their turn to be determined by the local field. The field acts to concentrate the distribution of dipoles around itself as closely as possible, in opposition to disruptive influences due to thermal agitation. If $n_i(\omega)$ gives the probability for any one molecule on lattice i to be in an orientation ω, then the part of the polarization due to the permanent molecular moment, μ, is given by

$$P_{di}=(\mu/2v)\int \cos\theta\, n_i(\omega)d\omega, \quad i=1,2, \tag{3}$$

where v is the volume of a unit cell, and θ the angle between z-axis and the orientation ω.

The most probable form of $n_i(\omega)$ under a given external field E and a given temperature T is of course to be determined from the minimum condition for the free energy of the crystal.

The entropy S associated with a macroscopic state specified by a set of distribution functions $n_i(\omega)$ is clearly

$$S=-k\sum_{i=1}^{3}\int n_i(\omega)\log n_i(\omega)d\omega, \tag{4}$$

k being the Boltzmann constant.

The internal energy U is given by definition

$$U=-\sum_{i=1}^{3}\int F_i dP_i=-E\sum_{i=1}^{3}P_i-\frac{1}{2\epsilon_0}\sum_{i=1}^{3}\sum_{j=1}^{3}f_{ij}P_iP_j. \tag{5}$$

The polarizations are each composed of two terms, since every molecule has, besides the orientation polarization, also the electronic polarization as well as the ionic polarization induced by the local field. Namely,

$$P_i=P_{di}+P_{ai}, \quad i=1,2; \quad P_3=P_{a3}, \tag{6}$$

together with

$$P_{ai}=\epsilon_0\alpha_1F_i/2v, \ i=1,2; \quad P_{a3}=\epsilon_0\alpha_3F_3/v, \tag{7}$$

where α_1 and α_3 are the polarizabilities of the molecules and ions respectively, excluding those due to the permanent dipoles.

Substituting (6) and (7) into (1), we can eliminate all the P_{ai} terms and can express the polarizations P_i in terms of only $n_i(\omega)$ and E. Then Eq. (5) can be transformed into the form

$$U=-\epsilon_0\kappa E^2-hE\sum_{i=1}^{2}P_{di}-\frac{1}{2\epsilon_0}\sum_{i=1}^{2}\sum_{j=1}^{2}g_{ij}P_{di}P_{dj}, \tag{8}$$

where κ, h, and g_{ij} are simple but somewhat lengthy expressions involving α_i and f_{ij}. Their explicit forms,

[5] J. C. Slater, Phys. Rev. **78**, 748 (1950).
[6] S. Roberts, Phys. Rev. **81**, 865 (1951).
[7] L. W. McKeehan, Phys. Rev. **43**, 913 (1933).

however, will be omitted here, since only certain combinations of them are of importance, and if combined they become rather simple as seen later.

Following the procedure in Kirkwood's theory,[2] extremalization of the free energy F, i.e.,

$$F=U-TS, \tag{9}$$

by means of the calculus of variations, with respect to the functions $n_i(\omega)$, subject to the normalization restraint,

$$\int n_i(\omega)d\omega=1, \quad i=1, 2, \tag{10}$$

leads to the following set of the most probable distribution functions:

$$n_i(\omega)=\frac{x_i}{4\pi \sinh x_i}\exp(x_i\cos\theta), \quad i=1, 2. \tag{11}$$

The parameters x_i must satisfy the following simultaneous equations

$$\frac{x_i}{\sigma}=\frac{2v}{\mu}h-\epsilon_0 E+\sum_{j=1}^{2} g_{ij}L(x_j), \quad i=1, 2, \tag{12}$$

where $\sigma=\mu^2/\epsilon_0 2vkT$, and $L(x)$ is the Langevin function,

$$L(x)=\coth x-1/x=(x/3)-(x^3/45)+(2x^5/945)-\cdots(x\ll 1). \tag{13}$$

Substituting (11) into (3), we obtain as dipolar parts

$$P_{di}=(\mu/2v)L(x_i), \quad i=1, 2. \tag{14}$$

The total polarizations of respective sublattices can easily be written in terms of these Langevin functions (see the next section).

Our approximation is clearly just equivalent to that of the Weiss molecular field theory of ferromagnetism, but seems to fit the present problem more closely, because the dipole forces underlying the dielectric phenomena are essentially long range in nature.

III. ANTIFERROELECTRICITY AND FERROELECTRICITY

When $E=0$, Eqs. (12) are reduced to

$$\frac{x_i}{\sigma}=\sum_{j=1}^{2} g_{ij}L(x_j), \quad i=1, 2. \tag{15}$$

If nonzero roots of x_i, of minimum free energy, are obtainable from these simultaneous equations, then the sublattices would be spontaneously polarized. From the symmetry of the equations, it is easy to see that either $x_1=x_2$ or $x_1=-x_2$. These two solutions correspond respectively to ferroelectric and antiferroelectric states, for which the dipolar polarizations are respectively oriented parallel and antiparallel to each other.[8]

[8] H. Takahashi discussed for the first time the possibilities for both ferroelectric and antiferroelectric dipole arrangements in perovskite type crystals by means of the Lorentz field correction as used by Slater; he presented his theory at the Symposium on Dielectrics, held by the Physical Society of Japan at T.I.T., on April 1, 1950. Published by the Society in hectograph.

(i) Antiferroelectricity

Putting $x_1=x=-x_2$, Eqs. (15) are reduced to

$$x/\sigma=(g_{11}-g_{12})L(x). \tag{16}$$

As is well known in the Weiss theory of ferromagnetism, there exists a critical value of σ, equal to $3/(g_{11}-g_{12})$, below which Eq. (16) has no real solution different from zero, and above which a single nonzero positive real root x, of minimum free energy, exists. This situation leads to a phase change of the second kind with a transition temperature, T_a, given by

$$T_a=(\mu^2/6k\epsilon_0 v)(g_{11}-g_{12}). \tag{17}$$

Substituting (17) into (16), we get a formula defining x as an implicit function of the reduced temperature T/T_a, namely

$$3L(x)/x=T/T_a. \tag{18}$$

Now it is easy to show that

$$g_{11}-g_{12}=(f_{11}-f_{12})\Big/\left\{1-\frac{\alpha_1}{2v}(f_{11}-f_{12})\right\}^2. \tag{19}$$

The polarizations of the lattices are calculated to be

$$P_1=-P_2=\frac{\mu}{2v}L(x)\Big/\left\{1-\frac{\alpha_1}{2v}(f_{11}-f_{12})\right\}$$
$$P_3=0. \tag{20}$$

(ii) Ferroelectricity

Putting $x_1=x_2=x$, Eqs. (15) are reduced to

$$x/\sigma=(g_{11}+g_{12})L(x). \tag{21}$$

In a similar manner to that used in (i), we can get a critical temperature

$$T_f=(\mu^2/6k\epsilon_0 v)(g_{11}+g_{12}), \tag{22}$$

and instead of (18) a similar relation

$$3L(x)/x=T/T_f. \tag{23}$$

Now it is easy to show that

$$g_{11}+g_{12}=\frac{2}{3}\Big/\left\{1-\frac{1}{3}\left(\frac{\alpha_1}{v}+\frac{\alpha_3}{v}\right)\right\}^2. \tag{24}$$

The polarizations of the lattices are calculated to be given by

$$P_1=P_2=\frac{1-\frac{1}{3}(\alpha_3/v)}{1-\frac{1}{3}(\alpha_1/v+\alpha_3/v)}\frac{\mu}{2v}L(x),$$
$$P_3=\frac{\frac{2}{3}(\alpha_3/v)}{1-\frac{1}{3}(\alpha_1/v+\alpha_3/v)}\frac{\mu}{2v}L(x). \tag{25}$$

(iii) Specific Heat Anomaly

The internal energy U, Eq. (8), can be transformed by means of (14), (17), and (22), if $E=0$, to

$$U=-(3k/2v)T_a L^2(x), \tag{26}$$

where T_c should read either T_a or T_f according to the type of the crystal concerned. Hence we have at once as the specific heat associated with rotational configurations, for one mole of the crystal,

$$\frac{C_v}{R}=\frac{x^2L'(x)}{1-3(T_c/T)L'(x)}, \tag{27}$$

R being the gas constant. Equation (27) is just the same as the one Kirkwood obtained; it yields at the transition point a lambda-type peak with a discontinuity of $2.5R$ at the high temperature side. Besides, it tends to a finite value, R, as the temperature approaches 0°K.

Since the relation between C_v and T/T_c is common for both ferroelectric and antiferroelectric cases, it is quite impossible to discriminate between them by means of only caloric measurements.

(iv) Relative Stability of the Two Solutions

It is easy to write the free energy (9) in the following explicit form:

$$F=\frac{kT}{2v}\sum_{i=1}^{\frac{n}{2}}\left[\log\frac{x_i}{4\pi\sinh x_i}+\frac{x_i}{2}L(x_i)-\frac{h\mu E}{2kT}L(x_i)\right]. \tag{28}$$

When $E=0$, the last term drops out; the remaining quantity in the bracket decreases monotonically with increasing $1/T$, because the derivative $\partial(F/T)/\partial(1/T)=U$ is always negative as seen from (26). Also, x increases monotonically with increasing $1/T$ as seen from Eq. (18) or Eq. (23).

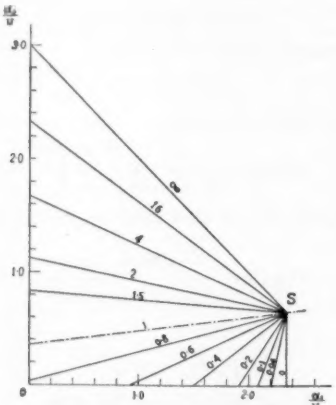


FIG. 2. The iso-Q lines: $Q=T_f/T_a$. All straight lines go through a singular point with the co-ordinates ($\alpha_1/v=2.3474$, $\alpha_2/v=0.6526$). In the region above the line $Q=1$, which is drawn as a chain line, the ferroelectric phase is more stable; in the region below it the antiferroelectric phase is more stable.

Suppose an arbitrary temperature T is given. Then a greater x, and accordingly a smaller free energy would be obtained if a case is chosen wherein the given temperature corresponds to a lower reduced temperature. That is to say, between ferroelectric and antiferroelectric cases, the one that has a higher critical temperature should be more stable.

By means of (17), (19), (22), and (24), the ratio Q of the two critical temperatures can be expressed in terms of the polarizabilities and the Lorentz factors;

$$Q=\frac{T_f}{T_a}=\frac{\frac{2}{3}}{\{1-\frac{1}{3}(\alpha_1/v+\alpha_2/v)\}^2}\Big/\frac{(f_{11}-f_{12})}{\{1-(\alpha_1/2v)(f_{11}-f_{12})\}^2}. \tag{29}$$

If an arbitrary value is given to Q, Eq. (29) defines a straight line in an $\alpha_1/v-\alpha_2/v$ plane, several of these lines are shown in Fig. 2. In the region below the special line for which $Q=1$, or

$$\alpha_2/v=0.3462+0.1304(\alpha_1/v), \tag{30}$$

the sublattices should be polarized in an antiferroelectric manner, while in the region above it they should be polarized in a ferroelectric manner. All lines go through a singular point S with the coordinates ($\alpha_1/v=2.3474$, $\alpha_2/v=0.6526$).

IV. DIELECTRIC CONSTANT

The electric displacement, D, namely

$$D=\epsilon_0E+\sum_{i=1}^{\frac{n}{2}}P_i, \tag{31}$$

can be written in terms of E and P_{di} by means of (1), (2), (6), and (7); i.e.,

$$D=\epsilon_0E+\frac{\alpha_1/v+\alpha_2/v}{1-\frac{1}{3}(\alpha_1/v+\alpha_2/v)}\epsilon_0E+\frac{1}{1-\frac{1}{3}(\alpha_1/v+\alpha_2/v)}\frac{\mu}{2v}\{L(x_1)+L(x_2)\}, \tag{32}$$

where x_i are to be obtained from Eqs. (12). The third term is not a linear function of x_i, and the x_i in their turn are not linear functions of the field strength E. Therefore, the dielectric constant may vary with a biasing dc field; this effect will be discussed in the next section. We shall here investigate only the initial value of the dielectric constant so defined that

$$[dD/d(\epsilon_0E)]_{E=0}=K. \tag{33}$$

The derivative of x_i with respect to E can be calculated from Eqs. (12), where the relations easily proved from direct calculation,

$$h=\tfrac{1}{3}(g_{11}+g_{12}), \tag{34}$$

may be utilized. Thus we get finally

$$K=1+\frac{\alpha_1/v+\alpha_2/v}{1-\frac{1}{3}(\alpha_1/v+\alpha_2/v)}+\frac{3}{1-\frac{1}{3}(\alpha_1/v+\alpha_2/v)}\frac{3(T_f/T)L'(x)}{1-3(T_f/T)L'(x)}, \quad (35)$$

where the value of x is to be obtained as a function of reduced temperature from (18) or (23) according to the type of stable solution. It is especially to be emphasized that the last term always contains T_f whether or not the crystal concerned is really ferroelectric.

$L'(x)/T$ tends to zero as the temperature approaches 0°K, so that the value of K at 0°K, K_0, consists of only first two terms;

$$K_0=1+\frac{\alpha_1/v+\alpha_2/v}{1-\frac{1}{3}(\alpha_1/v+\alpha_2/v)}, \quad (36)$$

or, if solved inversely with respect to the polarizabilities, we get the famous Clausius-Mosotti equation that

$$\frac{\alpha_1}{v}+\frac{\alpha_2}{v}=3\frac{K_0-1}{K_0+2}. \quad (37)$$

Substituting this into (35), we obtain

$$K=K_0+(K_0+2)\frac{3(T_f/T)L'(x)}{1-3(T_f/T)L'(x)}. \quad (38)$$

In a ferroelectric case, x is a function of T_f/T only, so that K is dependent only upon the reduced temperature in addition to K_0. Of course we are supposing that, even if there exist ferroelectric domains, the domain boundaries cannot move under such weak fields as are used in ordinary methods for measuring dielectric constants. Otherwise, Eq. (38) will not apply well to the observed values.

As T approaches T_f, $3(T_f/T)L'(x)$ tends to unity, so that the value of K goes to infinity at the transition point.

If $T>T_f$, then $L'(x)$ is always equal to $\frac{1}{3}$, so that we get

$$K=K_0+(K_0+2)T_f/(T-T_f). \quad (39)$$

Comparing this with the well-known Curie-Weiss law,

$$K=K_0+C/(T-T_p), \quad (40)$$

we see at once that the Curie constant C is

$$C=(K_0+2)T_f, \quad (41)$$

and that the paraelectric Curie point T_p is

$$T_p=T_f. \quad (42)$$

It turns out that if K_0 happens to be very large, the Curie constant C can be far larger than T_f. This is rather a remarkable result because it has been generally believed that, so far as the phase change concerned really arises from a system of rotatable dipoles, C must necessarily be of the same order of magnitude as T_f itself.[5]

In an antiferroelectric case, x is no longer a function of T_f/T but of T_a/T, so that when T approaches the critical temperature T_a, $3(T_f/T)L'(x)$ tends to not one but to $T_f/T_a=Q$. Consequently, the value of K at the Curie point, K_c, remains finite for this case, namely

$$K_c=K_0+(K_0+2)Q/(1-Q). \quad (43)$$

Or we can express the ratio Q in terms of the two special values on the $K-T$ curve, namely K_0 and K_c;

$$Q(=T_f/T_a)=(K_c-K_0)/(K_c+2). \quad (44)$$

Substituting this into (35), we can get a formula most convenient for practical calculation of the $K-T$ curve of a given antiferroelectric crystal,

$$K=K_0+(K_0+2)\frac{Q3(T_a/T)L'(x)}{1-Q3(T_a/T)L'(x)}. \quad (45)$$

If $T>T_a$, Eq. (45) leads to

$$K=K_0+(K_0+2)QT_a/(T-QT_a). \quad (46)$$

Comparing this with the Curie-Weiss law, we find that

$$C=(K_0+2)QT_a, \quad (47)$$

$$T_p=QT_a(=T_f). \quad (48)$$

From (47) we see that, if K_0 is very large and moreover if $K_c\gg K_0$, namely $Q\doteqdot 1$, C can again be very large compared with the transition temperature T_a. Since Q must be less than unity in order that an antiferroelectric state may be realized, the paraelectric Curie temperature for an antiferroelectric crystal should in general be lower than its transition temperature T_a.

Whereas the K-T curves for ferroelectric crystals are, roughly speaking, always similar to each other notwithstanding an arbitrary choice of the parameter K_0, and so especially if $K_0\gg 2$, the K-T curves for antiferroelectric crystals can show great variety of form for different choices of the parameters, K_0 and K_c. For examples, two extremely different forms are shown in Fig. 3. It may be seen that the curve (a) in the figure is of a form closely resembling that observed with some kinds of crystals known as containing rotatable polar molecules.

Besides, it is to be emphasized that the magnitude of the permanent electric moment of the molecules, μ, only influences the temperature scale as seen from Eqs. (17) and (22), but does not influence at all the value of dielectric constant itself. This is so entirely because the most important parameters, K_0 and K_c, are both completely determined by the polarizabilities of the crystal components. By means of Eqs. (29), (37), and (44) we can at once calculate the iso-K_0 lines as well as

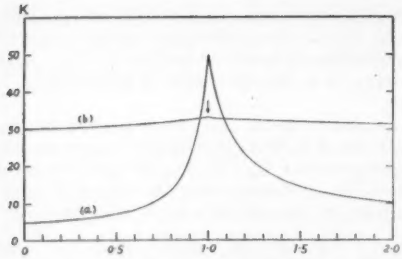


FIG. 3. Two examples of antiferroelectric $K-T$ curves; (a) for a crystal with $K_0=5$ and $K_e=50$; (b) for a crystal with $K_0=30$ and $K_e=33$.

the iso-K_e lines, the results being shown in Fig. 4. In the triangular region, above the line denoted by $K_e=$ infinity, for stable ferroelectric solutions, K_e is of course always infinity.

It may also be seen from Fig. 4 that, if both K_0 and K_e of a crystal happens to be very large, a point representative of the crystal should lie very near to the singular point S in the figure. In the neighborhood of S, the crystals may be very sensitive to any change, however slight, in the polarizabilities as well as in the cell volume. By way of illustration, we show in Table I the sets of polarizabilities calculated respectively for three specimens with very different peak values. The points representative of the three crystals described in Table I are thus situated within less than 0.5 percent from each other, so that they may be almost indiscernible if plotted in a figure of as small a size as Fig. 4. The $K-T$ curves of the last two examples are shown in Fig. 5.

It is also seen that, in a certain special case at least,

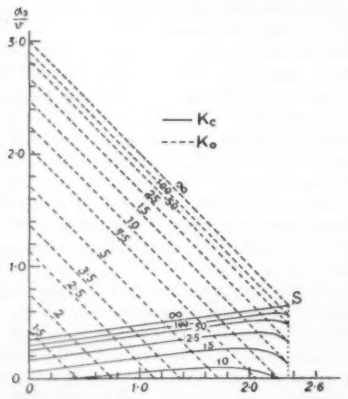


FIG. 4. The dashed lines represent the iso-K_0 lines; the solid lines represent the iso-K_e lines. In the triangular region above the straight line denoted by $K_e=$ infinity, i.e., the same line as that shown by a chain line in Fig. 2, K_e is everywhere infinity.

even the $K-T$ curve for an antiferroelectric crystal can attain an extraordinarily high value at its Curie point, though such a sharp peak is indeed unknown in antiferromagnetic cases. Therefore, if, on the other hand, K_e remains finite in a ferroelectric crystal as it actually does in many substances, it may happen that the entire $K-T$ curves resemble each other so much that it may be almost impossible to discriminate between ferroelectrics and antiferroelectrics from their $K-T$ curves alone.

V. RESPONSE TO A DC BIASING FIELD

In order to look for the points of contrast available to discriminate between ferroelectric and antiferroelectric crystals, we shall investigate their response to a dc biasing field. As seen in Sec. IV, the electric displacement (32) contains a nonlinear term composed of the Langevin functions, so that some variation of K should in general be expected under a biasing field.

TABLE I. Polarizabilities calculated for three specimens with different peak values.

Specimens			
K_0	K_e	α_1/ν	α_3/ν
100	1600	2.2719	0.6399
100	3200	2.2707	0.6411
100	infinity[a]	2.2695	0.6423

[a] The value tabulated here is the right lower end, corresponding to $Q=1$, of a line,

$$(\alpha_1/\nu)+(\alpha_3/\nu)=2.9118,\quad 0.6423\leq(\alpha_3/\nu)\leq 2.9118,$$

representing the general solutions for the condition $K_0=100$ and $K_e=$ infinity.

By means of the fundamental equations (12) together with (32), we get for an antiferroelectric case, to the second power in E,

$$\left[\sum_{i=1}^{2}\frac{dL(x_i)}{d\epsilon_0 E}\right]_E=2\left[L'(x)\frac{dx}{d\epsilon_0 E}\right]_0$$

$$+\left[\frac{L'''(x)\{1-3(T_a/T)L'(x)\}+9(T_a/T)\{L''(x)\}^2}{\{1-Q3(T_a/T)L'(x)\}\{1-3(T_a/T)L'(x)\}}\right]_0$$

$$\times\left(\frac{dx}{d\epsilon_0 E}\right)_0^2(\epsilon_0 E)^2, \quad (49)$$

where the suffixes outside the bracket denote respectively the values under E and zero field. The quantity x in the right-hand side is to be obtained from Eq. (18).

The first term contributes to the initial value of the dielectric constant, while the second term contributes to the change of K under a field E. By means of the series expansion (13) of the Langevin function, it is easy to prove that the second term is indeed positive, so that K should increase with increasing bias, though slowly at first since it is proportional to E^2.

On the other hand, for a ferroelectric case, we obtain

$$\left[\sum_{i=1}^{3}\frac{dL(x_i)}{d\epsilon_0 E}\right]_E = 2\left[L'(x)\frac{dx}{d\epsilon_0 E}\right]_0 + \left[\frac{2L''(x)}{1-3(T_f/T)L'(x)}\left(\frac{dx}{d\epsilon_0 E}\right)^2\right]_0 \epsilon_0 E. \quad (50)$$

As is easily seen, the second term is negative; accordingly K should decrease with increasing bias.

Hence the response to a biasing field, occurring in opposite directions, may be available as a convenient criterion for discerning between the two types of dielectric crystals, though of course these effects might be perceptible only in the immediate vicinity of the Curie point.

If the temperature is above the critical point, the difference between ferroelectric and antiferroelectric dipole configurations disappears. And there K decreases with increasing dc bias, as is seen from Fig. 7 in the next section.

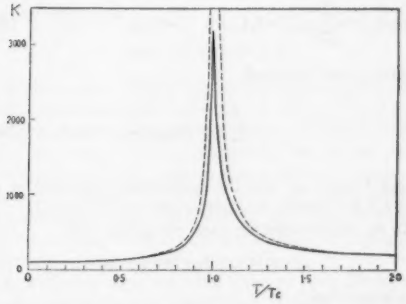


FIG. 5. The antiferroelectric and ferroelectric $K-T$ curves. The solid curve is for an antiferroelectric case with $K_0=100$ and $K_\infty=3200$; the dashed curve is for a ferroelectric case with $K_0=100$ and $K_\infty=$ infinity. T_c should read either T_a or T_f according to their nature.

VI. *D–E* CURVES AND A FORCED TRANSITION

Though the D–E curves for ferroelectric crystals are very familiar already, those for antiferroelectric crystals have not been discussed yet. We shall therefore take up this problem.

Taking a sum and difference of the fundamental Eqs. (12), we get at once the following set of equations convenient for the present purpose:

$$L(x_1)-\tfrac{1}{3}(T/T_a)x_1 = L(x_2)-\tfrac{1}{3}(T/T_a)x_2, \quad (51)$$

$$\tfrac{1}{3}(T/T_f)(x_1+x_2)-\{L(x_1)+L(x_2)\} = (6\nu/\mu)\epsilon_0 E. \quad (52)$$

From these, we can easily calculate the D–E curves for a given temperature T.

By way of illustration, we shall show the result at a temperature, say, $T=0.9837T_a$. In Fig. 6, the function

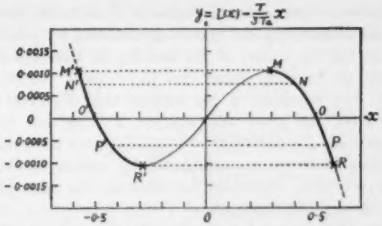


FIG. 6. The ordinate represents $L(x)-\frac{1}{3}(T/T_a)x$, for a temperature $T/T_a=0.9837$. The sets of points such as OO', PP', and so on satisfy Eq. (51). Antiferroelectric solutions can exist only between two horizontal lines MM' and RR'.

$L(x)-xT/3T_a$ is shown plotted against x. Equation (51) requires that each set of the roots x_i must always lie on a horizontal line, so that we get a series of the sets of points such as NN', OO', PP', and so forth. If an arbitrary set among these is substituted into Eq. (52), then we can find a field strength E which is necessary to maintain the specified set as stable roots. Besides, from Eq. (14) we can get the value of dipolar polarizations for this set of roots. The results thus calculated are shown plotted in Fig. 7.

Two characteristic points may be noted: The D–E curve show a slight upward curvature, in contradistinction to the well established downward curvature in ferroelectric D–E curves. The antiferroelectric D–E curve can exist only between the two limits, M and R, and if the field should exceed this threshold value the crystal would be forced to make a transition to another type of phase; of course this forced transition would be observed only when the crystal happens to be strong enough to bear a sufficiently intense field.

Whether this other phase to which the crystal is to make a forced transition is ferroelectric or paraelectric, depends entirely upon whether the given temperature T

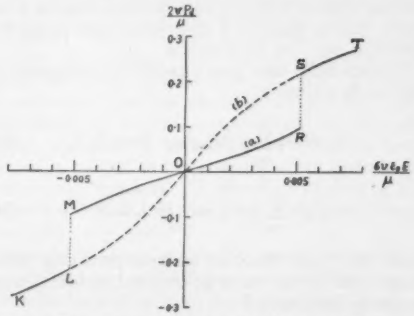


FIG. 7. The ordinate represents $L(x_1)+L(x_2)$, which is proportional to net dipolar polarization (see Eq. (14)) (a) for antiferroelectric phase at $T/T_a=0.9837$, (b) for paraelectric phase at $T/T_f=1.016$. If for a crystal $K_0=100$ and $K_\infty=3200$, then these two reduced temperatures correspond to the same actual temperature, and the P_d-E curve should go successively through the points K–L–M–O–R–S–T.

is higher or lower than the value of T_f expected for the crystal concerned, and this is determined by what we choose for the values of K_0 and K_c, so that the D–E curves of another phase can have great variety of form. For example, if we assume that $K_0=100$ and $K_c=3200$, the given temperature T turns out to be higher than T_f, namely $T=1.0161T_f$, since the ratio of the two Curie points calculated by means of (44) is $T_f/T_a=0.9681$. Therefore, for this case, the other phase must be a paraelectric one.

The D–E curve for this paraelectric phase is also obtainable from Eq. (52), in which $x_1=x_2=x$ must be used. The result is shown plotted in the same figure.

Though it is certain that an antiferroelectric phase cannot persist outside of the limits M and R, yet these extreme points must not be regarded as representing true transition points between the two phases. It is in general expected from a thermodynamical point of view that the free energy curves for both phases will intersect at some points lying somewhere inside the extremities. However, after calculation in the present case, it has turned out that both the free energy curves seem, within the accuacy of our calculation, to intersect at the very ends of the antiferroelectric range, namely, just at M and R. We have shown this situation by the two vertical dotted lines in Fig. 7.

Besides, we have a proof for the statement given at the end of Sec. V that, for the paraelectric case, K should decrease with increasing dc bias. Indeed the curve (b) in Fig. 7 is clearly seen to have a remarkable downward curvature.

VII. TRANSVERSE APPLICATION OF FIELD

When the sublattices are spontaneously polarized in the direction of the z-axis, the crystal is no longer of cubic symmetry but is of tetragonal symmetry, so that the dielectric constant tensor will be completely determined if we know the dielectric constant in an arbitrary direction perpendicular to z-axis. We shall examine what results will be obtained if an external field is applied parallel to, say, the $+x$ direction.

The local fields then have x as well as z components; they can be written as

$$F_{iz}=E+\sum_{j=1}^{3} f_{ij}P_{jz}/\epsilon_0, \quad i=1, 2, 3, \tag{53}$$

$$F_{ix}=\sum_{j=1}^{3} f_{ij}P_{jx}/\epsilon_0, \quad i=1, 2, 3 \tag{54}$$

where the f_{ij} are the same as those previously introduced, while the f_{ij}^x are to be obtained, as before, from a paper by McKeehan;[7]

$$\begin{aligned} &f_{ij}^x=f_{ji}^x,\\ &f_{11}^x=f_{22}^x=\tfrac{1}{3}-0.2130,\\ &f_{12}^x=\tfrac{1}{3}+0.2130,\\ &f_{13}^x=f_{23}^x=f_{33}^x=\tfrac{1}{3}. \end{aligned} \tag{55}$$

The polarizations are

$$\begin{aligned} &P_{ai2}=\epsilon_0\alpha_1F_{i2}/2v, \quad P_{ai3}=\epsilon_0\alpha_1F_{i3}/2v; \quad i=1, 2\\ &P_{a32}=\epsilon_0\alpha_3F_{32}/v, \quad P_{a33}=\epsilon_0\alpha_3F_{33}/v, \end{aligned} \tag{56}$$

and

$$\begin{aligned} &P_{di2}=(\mu/2v)\int \sin\theta\cos\varphi\; n_i(\omega)d\omega,\\ &P_{di3}=(\mu/2v)\int \cos\theta\, n_i(\omega)d\omega, \quad i=1, 2. \end{aligned} \tag{57}$$

Proceeding in a similar manner to that used in Sec. II, we get the most probable distribution functions,

$$\begin{aligned} n_i(\omega)=\frac{x_i}{4\pi\sinh x_i}\exp[x_i(&\cos\theta_i\cos\theta\\ &+\sin\theta_i\sin\theta\cos\varphi)], \quad i=1, 2, \end{aligned} \tag{58}$$

where the four parameters x_i and θ_i are to be solved as the relevant roots of the following set of equations:

$$x_i\cos\theta_i/\sigma=\sum_{j=1}^{2} g_{ij}\cos\theta_jL(x_j), \quad i=1, 2, \tag{59}$$

$$\begin{aligned} x_i\sin\theta_i/\sigma=h^x(2v/\mu)\epsilon_0E&\\ +\sum_{j=1}^{2} g_{ij}^x\sin\theta_jL(x_j), \quad i=1, 2,& \end{aligned} \tag{60}$$

where h^x and g_{ij}^x are the coefficients composed of α_i and f_{ij}^x in a similar manner to that used in Sec. II.

In an antiferroelectric case, we get at once

$$x_1=x_2=x, \quad \theta_1=\Theta=\pi-\theta_2,$$

and accordingly

$$x/\sigma=(g_{11}-g_{12})L(x), \tag{61}$$

$$\sin\Theta=\{h^x(2v/\mu)\epsilon_0E\}/\{[(g_{11}-g_{12})-(g_{11}^x+g_{12}^x)]L(x)\}. \tag{62}$$

We see from (61) that the $x-T/T_a$ relation is not influenced at all by a transverse application of any external field whatever. The dipolar polarizations calculated by means of (57) and (58) turn out to be, for this solution.

$$\begin{aligned} &P_{d1x}=P_{d2x}=(\mu/2v)L(x)\sin\Theta,\\ &P_{d1z}=-P_{d2z}=(\mu/2v)L(x)\cos\Theta. \end{aligned} \tag{63}$$

This tells us that the axes of cylindrical distribution of the dipole orientations are both rotated away from their original directions toward the $+x$ direction by the angle Θ given in (62).

Substituting these polarizations into the electric displacement,

$$D_z=\epsilon_0E+\sum_{i=1}^{3} P_{iz},$$

we obtain

$$D_s=\left[1+\frac{(\alpha_1/v)+(\alpha_3/v)}{1-\frac{1}{2}(\alpha_1/v+\alpha_3/v)}+\frac{1}{1-\frac{1}{2}(\alpha_1/v+\alpha_3/v)}\times\frac{h^2}{(g_{11}-g_{12})-(g_{11}^2+g_{12}^2)}\right]\epsilon_0 E. \quad (64)$$

We see at once that K_s is independent of temperatures, because the terms in the bracket are all composed of only constant quantities. And since it is easily proved that

$$g_{11}^2+g_{12}^2=g_{11}+g_{12}, \quad (65)$$

$$h^2=h \quad (66)$$

this constant value of K_s turns out to be just equal to the value of K_c previously found for the dielectric constant at the Curie point.

In the ferroelectric case, however, we find a rather undesirable result; K_s turns out to be infinite at all temperatures below the critical point. But this result is consistent with the fact that K_c is really infinite for ferroelectric cases.

VIII. DISCUSSIONS

So far we have shown that a CsCl type crystal containing rotatable polar molecules can show either ferroelectric or antiferroelectric properties depending solely upon the values of the polarizabilities of its components. In addition, we have examined in what respects the two phases will be able to show contrasts of phenomenological importance, and in what respects they will behave very similarly notwithstanding the essential difference in their dipole configurations.

As is clearly seen, we have hitherto been discussing merely the theoretical possibilities to be derived from our hypothetical model, though of course with the hope that the results obtained would be of some use in interpreting the dielectric properties of some actual crystals. Now let us examine what modification must be introduced when some other model is considered.

Careful inspection of the present theory tells us at once that, so long as a crystal really contains rotatable dipoles and so long as a decomposition of the dipole lattice into a set of two sublattices is allowable from the first as in the present model, the general aspects of the theoretical predictions should remain valid however complicated the crystal structure may be. Of course, then the factors such as h and g_{ij} will become more complicated and some of the results such as shown in Figs. 2 and 4 may be changed very considerably, while some other predictions such as those presented in Eqs. (18), (23), (27), (38), and also in Eqs. (49) and (50) will be valid in just the same forms.

Also, even when a crystal contains dipoles due to the finite shift of ions on a specified line instead of freely rotatable polar molecules, as perhaps is the case with some actual crystals like rochelle salt, most of the fundamental equations will be retained, except that Eqs. (3) and (4) should be replaced by

$$P_{di}=(\mu/2v)[n_i(+)-n_i(-)], \quad (67)$$

$$S=k\sum_{i=1}^{z}\log\frac{[n_i(+)+n_i(-)]!}{n_i(+)!n_i(-)!}, \quad (68)$$

where $n_i(+)$ and $n_i(-)$ denote the number of the dipoles on lattice i in the directions $+z$ and $-z$, respectively. Then all the Langevin functions, wherever they may appear, will be replaced by another similar function $\tanh x$, or $L_1(x)$, as is at once expected from an analogy to the Weiss-Heisenberg theory of scalar spin ferromagnetism and antiferromagnetism.[9] Most of the theoretical predictions, except those concerning the form of the specific heat curve, will, however, again remain valid, at least qualitatively.

Since last year we have been investigating the dielectric properties of ceramic $PbZrO_3$. Its dielectric constant-temperature curve resembles that of the well-known ferroelectric $BaTiO_3$ so much that it has been suggested that $PbZrO_3$ might also be a kind of ferroelectric.[10]

After our detailed investigation, however, we have found:[11]

(i) Its dielectric constant increases slightly with increasing dc bias below the Curie point, while it decreases with increasing bias above the Curie point.

(ii) Its D–E curves observed on a cathode-ray oscilloscope show a slight upward curvature just below the Curie point, and moreover when the amplitude of the applied alternating field is greater than 30 kv/cm at 226°C, an impressive form of queer D–E curve is seen, very similar to that shown in Fig. 7, with abrupt increases of polarization on both ends.

We inferred, therefore, that this material might rather be an antiferroelectric, though its actual structure was unknown at that time. In the meantime, this surmise has proved to be true. Indeed, very recently, Sawaguchi, Maniwa, and Hoshino[12] have succeeded in proving by means of an x-ray analysis of single crystals of $PbZrO_3$ the existence of an essentially antiferroelectric dipole arrangement in a plane perpendicular to its c-axis, though the question whether it can be slightly ferroelectric in c-direction has been left open. Roberts has subsequently measured the piezoelectric effect of this material and has found that it certainly exists, though it is in fact very small.[13]

Considering these events, it seems reasonable to suggest that most of the theoretical predictions may be valid also for more general types of crystals, though of

[9] Y. Takagi, Proc. Phys.-Math. Soc. Japan **23**, 553 (1941); **24**, 333 (1942).
[10] S. Roberts, J. Am. Ceram. Soc. **33**, 63 (1950).
[11] Shirane, Sawaguchi, and Takagi, Phys. Rev. **84**, 476 (1951).
[12] Sawaguchi, Maniwa, and Hoshino, Phys. Rev. **83**, 1078 (1951).
[13] S. Roberts, Phys. Rev. **83**, 1078 (1951).

course qualitatively, even when their accurate structures are not known initially.

The most undesirable prediction derived from our theory may probably be that concerning the transverse dielectric constant; that is, K_x is far larger than is reasonable. But it is certain that this defect has arisen chiefly because no crystal anisotropy whatever is taken into account. If a crystal is spontaneously polarized in any one direction, there is no reason at all why it must still be cubic. In general, the crystal would preferably become strained, since the strains would stabilize all the more the specified axis of dipole orientation. Any amount of rotation of the axis of polarization away from the easiest direction would be realizable only when the transverse field is strong enough to overcome the potential barriers surrounding the easiest axis; this effect will certainly diminish the K_x value considerably.

In this respect, it seems very interesting to notice that in single crystals of $BaTiO_3$, the dielectric constant in the direction perpendicular to its spontaneous polarization is found to be far larger than the dielectric constant measured in the parallel field.[14]

I wish to express my hearty thanks to the staff of our laboratory in the Tokyo Institute of Technology, especially to Messrs. T. Oguchi, G. Shirane, and E. Sawaguchi for their very helpful discussions.

[14] W. Merz, Phys. Rev. **75**, 687 (1949).

PHYSICAL REVIEW VOLUME 85, NUMBER 2 JANUARY 15, 1952

Interaction between the *d*-Shells in the Transition Metals. IV. The Intrinsic Antiferromagnetic Character of Iron

C. ZENER*
Westinghouse Research Laboratories, East Pittsburgh, Pennsylvania
(Received September 19, 1951)

The principles developed in the previous papers of this series are applied to interpret the presence of ferromagnetism in b.c.c. iron, its absence in f.c.c. iron. It is found necessary to introduce a new principle, namely that for minimum energy the *d*-shell electrons are distributed among the atoms so as to maximize the total number of pairs of electrons having like spin within the individual atoms. This principle leads to the viewpoint that, at least on a local scale, the b.c.c. iron lattice consists of the superposition of two interpenetrating simple cubic lattices of atoms having *d*-shell spins of different magnitudes and of opposite direction. The same principles lead directly to the observed saturation magnetizations of all iron base alloys.

I. INTRODUCTION

IN the previous papers[1] of this series the author has proposed that for the transition elements the magnetic structure of the *d*-shells of the individual atoms is essentially the same in the metallic as in the gaseous state, that ferromagnetic coupling between the incomplete *d*-shells in metals arises from an indirect coupling via the conduction electrons, and that the direct interaction between these shells is antiferromagnetic. These proposals have provided an unforced interpretation of the occurrence of ferromagnetism in the periodic table, of the occurrence of b.c.c. lattices in the transition metals, and have provided a quantitative calculation of the Weiss coupling factor. The essential validity of these proposals has been further strengthened by the observation of the antiferromagnetic behavior of the magnetic susceptibility of chromium by McGuire and Kriessman,[2] by the observation of the anomalously low value of the electronic specific heat of chromium at low temperatures by Friedberg,[3] by the computation of the binding energy of tungsten by Hsu,[4] by the analysis of the elastic constants of tungsten by Isenberg,[5] by the analysis of the Heusler alloys by Heikes,[6] and by the detailed analysis of the magnetic properties of alloys by Carr[7] in a forthcoming paper. Slater[8] has correlated the more usual band viewpoint with the atomistic viewpoint adopted in this series. In the second paper of this series it has been pointed out that, in the absence of *s* electrons, electrical conductivity by *d* electrons can also introduce ferromagnetic coupling. This relation between electrical conductivity by *d* electrons and ferromagnetic coupling has recently received further support from observations upon the ferrites. Below their Curie temperature a second critical temperature, or narrow temperature range, is found in which both the magnetic properties and the electrical conductivity suffer large abrupt changes. (Manganese ferrite,[9] iron

* Associate Director, Westinghouse Research Laboratories East Pittsburgh, Pennsylvania.
[1] C. Zener, Phys. Rev. **81**, 440 (1951); **82**, 403 (1951); **83**, 299 (1951).
[2] T. R. McGuire and C. J. Kriessman, Jr., Phys. Rev. **82**, 774 (1951).
[3] S. Friedberg, Phys. Rev. (to be published).
[4] Yee-Chuang Hsu, Phys. Rev. **83**, 975 (1951).
[5] I. Isenberg, Phys. Rev. **83**, 637 (1951).
[6] R. Heikes, Phys. Rev. **84**, 376 (1951).
[7] W. J. Carr, Phys. Rev. (to be published).
[8] J. C. Slater, Phys. Rev. **82**, 538 (1951).
[9] T. Okamura and J. Simoizaka, Phys. Rev. **73**, 664 (1951).

ferrite,[10,11] nickel ferrite,[12] cobalt ferrite.[13]) An outstanding problem remains unsolved, namely, the structure sensitivity of the magnetic properties of iron. The b.c.c. phase of iron is ferromagnetic, the f.c.c. phase is not ferromagnetic. It is the purpose of the present paper to provide an understanding of this structure sensitivity.

The nonoccurrence of ferromagnetic coupling in f.c.c. iron cannot be dismissed merely by the observation that the transformation temperature from b.c.c. to f.c.c. iron lies above the Curie temperature of the former. If in the f.c.c. phase of the Fe−Co and of the Fe−Ni systems[14] one attempts to extrapolate the Curie temperature to the pure f.c.c. Fe phase, one finds this Curie temperature to lie at best in the vicinity of the absolute zero. Strong evidence of the absence of ferromagnetic coupling in the f.c.c. phase of iron is obtained from magnetic susceptibility measurements.[15]

An interpretation of the difference in the magnetic properties of the two phases of iron has been given by Néel[16] in terms of the original Heisenberg theory. In b.c.c. iron the nearest and next nearest neighbors are at distances of 2.48 and 2.86A, respectively. In f.c.c. iron corresponding neighbors are at distances of 2.52 and 3.58A, respectively. According to Néel, of these interatomic distances only 2.86A for next nearest neighbors in the b.c.c. phase lies within the appropriate range for ferromagnetic coupling. This interpretation suffers from the same disability as does the original Heisenberg theory, namely the lack of a theoretical basis for the existence of a range of interatomic distances within which the direct exchange interaction has the appropriate sign and magnitude for ferromagnetic coupling.

As one passes to the left along the first transition period from Ni through Co, Fe, Mn to Cr and V, the d-shell gradually expands. According to the previously developed viewpoint, the direct interaction between adjacent incomplete d-shells always tends to an antiferromagnetic alignment of spins. In Ni and Co, where the overlapping of adjacent d-shells is only very slight, the ferromagnetic indirect coupling via the conduction electrons dominates, leading to ferromagnetism. By the time Cr and V are reached, the overlap of adjacent d-shells is sufficiently great so that the direct antiferromagnetic coupling dominates, thereby stabilizing the b.c.c. lattice with respect to the f.c.c. lattice. Fe and Mn lie in the transition region between ferromagnetism and antiferromagnetism. The absence of ferromagnetism in f.c.c. iron leads us to suspect that a straightforward quantitative application of the above viewpoint would lead to the absence of ferromagnetism in all phases of iron, but that the b.c.c. phase possesses an intrinsic flexibility which, when properly considered, leads to ferromagnetism.

It is well recognized that exchange energy plays a dominant role in magnetism. The necessity for maximizing the negative exchange energy between electrons in the same d-shell assures us that the electrons in each incomplete d-shell will be distributed in the various quantum states so as to give the maximum possible net magnetic moment consistent with Pauli's exclusion principle. The negative exchange energy between the d-shell electrons and the conduction electrons gives rise to the indirect coupling between d-shells that leads to ferromagnetism. The positive exchange energy between electrons in adjacent d-shells gives the antiferromagnetic coupling that stabilizes the b.c.c. lattice of the transition metals in the V and Cr columns. The negative exchange energy between the conduction electrons themselves just about halves the coupling factor γ which appears in the author's formulation[1] of the theory of ferromagnetism. In previous discussions of ferromagnetism no attempt has been made further to increase the total negative exchange energy between electrons in the same d-shell by a transfer of d electrons from one atom to an adjacent atom. Simple considerations show that the total negative exchange energy will be increased by such a transfer if the atoms contain incomplete half d-shells.

As an example, we shall consider the particular case of iron. The outer configuration of the isolated Fe atoms is $3d^64s^2$. The necessity for minimizing the Fermi kinetic energy of the conduction electrons assures us that at least one $4s$ electron will be demoted into the $3d$ shell when the atoms condense to form the solid phase. The electronic configuration more appropriate to the metal is therefore $3d^74s$. In order to maximize the negative exchange energy between the electrons in the d-shell, we must place 5 electrons in one half-shell (Hund's rule). This distribution we shall symbolize as $(+++++, --)$. We now investigate the change in the negative exchange energy associated with the transfer of two electrons from the d-shell of one Fe atom to the d-shell of a second Fe atom. Such an exchange may be represented as

$$2(+++++, --) \rightarrow (+++++,) + (+++++, ----). \quad (1)$$

Now the exchange energy within a given half-shell is proportional to the number of pairs which may be formed between the electrons in the half-shell. The exchange energy in the positive half-shells is unaltered by such a transfer. On the other hand, the total number of pairs in the negative half-shells is increased from 2 to 6. The negative exchange energy is therefore increased by the transfer (1).

[10] C. A. Domenicali, Phys. Rev. **78**, 458 (1951).
[11] H. Watanabe and N. Tsuya, Sci. Repts. Tohoku Univ. **2**, 29 (1950).
[12] T. Okamura, Nature **168**, 162 (1951).
[13] C. Guillaud and H. Creveaux, Compt. rend. **230**, 1256 (1950).
[14] M. Hansen, *Aufbau der Zweistofflegierungen* (Edward Brothers, Inc., Ann Arbor, 1943), pp. 486, 697.
[15] R. Becker and W. Döring, *Ferromagnetismus* (Edward Brothers, Inc., Ann Arbor, 1943), p. 27.
[16] L. Néel, Ann. Phys. **5**, 270 (1936).

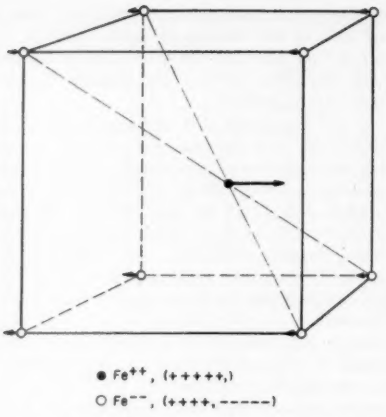


FIG. 1. Postulated spin structure of b.c.c. iron.

In the b.c.c. phase of iron, a maximizing of the total negative exchange energy within the *d*-shells will lead to the configuration shown in Fig. 1. The arrangement of the spins on a regular Cs Cl type lattice will minimize the electrostatic energy. The antiparallel arrangement of spins of the two simple cubic lattices leads to the observed saturation magnetic moment per atom. The saturation magnetic moment per atom of the arrangement of Fig. 1 is 2.0 μ_B, the observed value is 2.2 μ_B. The remaining 0.2 μ_B can arise from the polarization of the conduction electrons. This antiferromagnetic arrangement is consistent with the view that in iron the direct interaction between the *d*-shells and nearest neighbors is sufficiently large to overcome the indirect ferromagnetic coupling via the conduction electrons. The above considerations give a ready interpretation as to why the b.c.c. phase but not the f.c.c. phase of iron is ferromagnetic. No ordered arrangement of the charged atoms on a f.c.c. lattice will lower the electrostatic energy as much as on a b.c.c. lattice.

Two apparently serious objections might be raised against the above developed theory that the b.c.c. lattice is intrinsically polar in nature. The first objection is that the gain in exchange energy upon forming such a polar lattice will be more than offset by the rise in electrostatic energy. Upon taking the value of the exchange integrals computed by Slater[17] for nickel, we find the gain in exchange energy for a singly ionized lattice to be 0.9 ev per ion pair. In Sec. II we attempt to estimate the electrostatic energy. We find that it is not unreasonable for this electrostatic energy to be less than the above gain in exchange energy. The second objection is that such polarization should be detected by neutron diffraction, while Dr. Shull has informed the author that neutron diffraction reveals no such anomalies as would be anticipated from an ordered polar lattice in which the given atoms maintain the same polarity in a time at least as long as is required for a neutron to traverse a domain.† These negative results still allow the possibility of a polar lattice in which the polarity of a given atom fluctuates rapidly. Empirical evidence for a transfer of charge between atoms in alloys containing first row transition metals has been presented by Vogt,[18] for a transfer of charge between Mn atoms in Mn_2Sb by Guillaud,[19] and for a transfer of charge between Mn^{+++} ions in Mn_3O_4 by Verwey.[20]

[17] J. Slater, Phys. Rev. **49**, 537 (1936).

II. ENERGETICS OF FORMATION OF POLAR LATTICES

In order to gain a physical insight into the factors determining the work required to polarize a b.c.c. lattice of like atoms, we shall first consider the case of a lattice of metal spheres. When the spheres are isolated, the work expended in removing a charge q from one sphere of radius R is $q^2/2R$, and a like work is expended upon placing this charge upon a similar second sphere. The work of polarizing the isolated atoms is therefore q^2/R per pair. When we now bring the polarized atoms together to form a CsCl type lattice, we gain back a certain amount of work. If we keep the charge on each sphere uniformly distributed, we gain back the energy 1.76 q^2/D, where D is the closest distance between oppositely charged spheres. Upon taking the spheres to be close packed, i.e., upon taking D to be $2R$, we find that 88 percent of the originally expended energy is gained back. The remaining 12 percent of the originally expended energy corresponds, in the case of spheres having the size of iron atoms, to 1.4 ev for q equal to the charge of one electron.

Essentially the same relation between net work and work of polarization is found for all atoms for which data are available. Thus in the case of the halogens the work of polarization is the ionization potential minus the electron affinity. By taking D as the sum of the radii of the halogen negative ion and of the positive ion of the next succeeding alkali metal, we obtain the data of Table I.

From the above examples, it appears that it would require less than 2 ev per atom pair to transfer an electron from the *d*-shell of each iron atom of one simple cubic lattice to the *d*-shell of each iron atom of the second simple cubic lattice, a redistribution of the conduction electrons being restrained. If we now allow the conduction electrons to be redistributed to minimize the energy, it appears reasonable that the net work will be lowered below the 0.9 ev per atom pair gained in the increase in the negative exchange energy.

† *Note added in proof:* The most recent experiments of Dr. Shull do indeed indicate an antiferromagnetic structure in Cr.

[18] E. Vogt, Z. Metallk. **42**, 155 (1951).

[19] G. Guillaud, J. phys. et radium **12**, 239 (1951).

[20] Verwey, Braun, Gorter, Romeijn, and van Santen, Z. physik. Chem. (to be published).

III. MAGNETIC STRUCTURES OF IRON BASE ALLOYS

The variation of the saturation magnetic moment in the b.c.c. alloys of iron with other first row transition metals is given by the rough rule that[21]

$$d\mu/dN = \mu_B, \tag{2}$$

where μ is the saturation magnetic moment per atom, and N is the average number of electrons. Once the average magnetic moment has reached 2.5 μ_B, further additions result in a reversal in sign of $d\mu/dN$. This rule may be most readily understood by tentatively adopting the viewpoint that equivalent lattice sites are occupied by the same number of electrons, irrespective of nuclear charge. Such a viewpoint is in accord with the strict application of the band model. The addition of atoms of atomic number less than that of iron results in a decrease in the average number of electrons. Since the positive ions have a complete half-shell, this decrease in number of electrons will occur solely in the negative ions, thereby resulting in an increase in the magnetic moment of the negative ions, in a decrease of magnetic moment of the negative ions, in a decrease the over-all moment by the amount given by (2). Similarly, the addition of atoms with atomic number greater than that of iron results in a decrease in the magnetic moment of the negative ions, and hence in an increase in the over-all magnetic moment by the amount given by Eq. (2). Once the negative ions have completed their second half-shells, further electrons must be accommodated by the positive ions. Since these additional electrons must now start a new half-shell, their spins will be opposed to those already on the positive ions, and they hence contribute to a decrease in the average magnetic moment. The maximum value (5/2) μ_B of the average magnetic moment thereby receives a natural interpretation.

Adherence to the above strict band model leads, however, to unreasonably large charges on the ions, as indicated by the second and third columns of Table II. Thus in Fe—V alloys vanadium atoms on the negative ion sites will have an excess of 5 electrons in the d shell. It is now reasonable to expect that most of these excess electrons will be redistributed on the neighboring negative iron ions. Such a redistribution of electrons among the negative ion sites, all of which have incomplete d shells with magnetic moments pointing in the same direction, will not change the over-all magnetic moment. The partial relaxation of unreasonably large negative charges will therefore not change the average magnetic moment, and hence leaves unaltered rule (2). In the case of Fe—Ni alloys, however, we must relax the unreasonably large charge on the positive Ni ions. This relaxation will be accomplished by a transfer of electrons from the Fe^{--} ions to the Ni^{++++} ions. These electrons will suffer a reversal of spin in this transfer, thereby resulting in a change in the over-all magnetic moment. A relaxation from a quadruply charged positive Ni ion to a doubly charged positive Ni ion will in fact precisely cancel the increase in average magnetic moment of 2 μ_B per nickel atom predicted by the strict band model. This expected deviation of the Fe—Ni alloys from rule (2) is precisely that found experimentally.

[21] R. M. Bozorth, *Ferromagnetism* (D. Van Nostrand Company, Inc., New York, 1951), Fig. 10-11.

TABLE I. Energy of formation of polar lattice of halogen ions (energy in ev).

	Work of polarization	Madelung energy	Net work
Cl	9.0	8.0	1.0
Br	8.1	7.3	0.8
I	7.1	6.5	0.6

It is anticipated that the triple charge of Co^{+++} will relax at most only to a slight extent, since relaxation to Co^{++} would leave one d electron alone in a half-shell. That a slight relaxation does occur, with a concomitant slight reduction in the over-all magnetic moment, is indicated by the fact that ordering in a 50–50 Fe—Co alloy increases the average magnetic moment by about 4 percent.[22] In such an ordered alloy the iron atoms will occupy all the positive sites, the cobalt atoms all the negative sites.

Only the experiments upon the Fe—Mn alloys[23] are in conflict with the above theory. Observations up to 4 percent Mn indicate a variation of μ with Mn concentration about twice that given by (2). In view of the difficulty of being certain that a single phase is obtained in such alloys, the author is inclined to blame this disagreement upon the experiments rather than upon the theory.

In pure b.c.c. iron it is to be expected that the polarity of a given iron atom will rapidly fluctuate. Such fluctuations will however be eliminated in an ordered alloy when the alloying atoms have a distinct preference for one sign of polarity. Such will be the case, for example, in Fe_3Si and Fe_3Al, the alloying atoms here preferring to occupy the site of an Fe^{++} ion rather than that of an Fe^{--} ion. As indicated in Fig. 2, the negative Fe ions will occupy the corners of a simple cubic lattice. Half the centers will be filled by positive

TABLE II. Magnetic structure of disordered iron base alloys.

Alloy	Strict band model →Site	Strict band model ←Site	Realistic model →Site	Realistic model ←Site
Fe—V	V^-	V^{----}	V^-	V^-
Fe—Cr	Cr	Cr^{----}	Cr	Cr
Fe—Mn	Mn^+	Mn^{---}	Mn^+	Mn^{--}
Fe	Fe^{++}	Fe^{--}	Fe^{++}	Fe^{--}
Fe—Co	Co^{+++}	Co^-	Co^{+++}	Co^{--}
Fe—Ni	Ni^{++++}	Ni	Ni^{++}	Ni^-

[22] J. E. Goldman and R. Smoluchowski, Phys. Rev. **75**, 310 (1949).

[23] C. Sadron, Ann. Phys. **17**, 407 (1932).

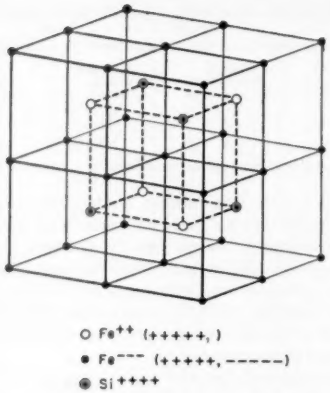


FIG. 2. Postulated spin structure of Fe_3Si.

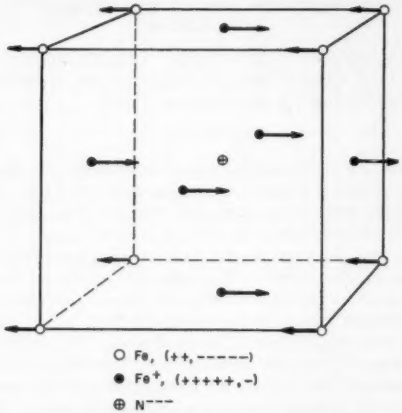


FIG. 3. Postulated spin structure of Fe_4N.

Fe ions, half by positive Si or Al ions.[24] In the case of Fe_3Si, Si has just enough valence electrons to fill completely the incompleted half d shells of the surrounding negative Fe ions. Such complete ionization of the Si atoms would give rise to the spin structure postulated in Fig. 2. Here only the Fe^{++} ions have a magnetic moment, of value 5 μ_B. The computed average magnetic moment is thus (5/4) μ_B per atom, in agreement with experiment.[25] The average magnetic moment of Fe_3Al has essentially the same value.[26] A more realistic computation would leave the negative iron ions with a slightly incomplete half-shell, the resulting decrease in the contribution of the d shells to the net magnetic moment being compensated by the magnetic polarization of the conduction electrons.

We shall finally analyze the ferromagnetism in f.c.c. iron induced by the solution of nitrogen. Face-centered cubic iron becomes ferromagnetic[27] when the center of each unit cube is occupied by a nitrogen atom. In contrast to Al, we expect each nitrogen atom to absorb three electrons from the neighboring Fe atoms. We thereby obtain a structure in which Fe^+ ions occupy the faces, Fe atoms the corners of the unit cube. The removal of an electron from the d shell of an Fe atom results in a contraction of this shell, thereby reducing the direct antiferromagnetic interaction with the nearest neighbors having similarly contracted d shells. If we now assume that this reduction is sufficient to allow for a ferromagnetic coupling via the conduction electrons, while the interaction between the Fe^+ ions and the neutral Fe atoms at the corners is still antiferromagnetic, we arrive at the spin structure shown in Fig. 3. The average magnetic moment per atom computed from this spin structure is (9/4) μ_B, in agreement with the observed value of 2.22.[28]

[24] R. M. Bozorth, *Ferromagnetism* (D. Van Nostrand Company, Inc., New York, 1951), Fig. 7-2.
[25] See reference 24, Fig. 4-9.
[26] See reference 24, Fig. 7-8.
[27] See reference 24, Fig. 7-36.
[28] See reference 24, p. 240.

Theory of Antiferromagnetic Resonance

F. KEFFER AND C. KITTEL
Department of Physics, University of California, Berkeley, California
(Received October 1, 1951)

The spin resonance condition $\omega/\gamma=H_0\pm[H_A(2H_E+H_A)]^{\frac{1}{2}}$ previously given by Kittel for a disk-shaped single-domain uniaxial or cubic antiferromagnetic crystal at 0°K with H_0 parallel to the domain axis is extended by classical calculations to cover finite temperature, ellipsoidal shape, orthorhombic symmetry, generalized two-lattice anisotropy, and arbitrary static field direction. The normal precessional modes are discussed. A quantum-mechanical derivation of the resonance equations is carried out by the method developed by Van Vleck for ferromagnetic resonance; no new features are introduced by the quantum-mechanical calculation. Several factors contributing to the line width are considered. Existing experimental data on antiferromagnetic resonance are reviewed; the data are scanty and taken in circumstances not closely related to the situation envisaged by the theory.

I. INTRODUCTION

WE are concerned in this paper with the theory of electron spin resonance absorption in antiferromagnetic crystals, and we find results exhibiting very significant differences from the results with ferromagnetic and paramagnetic crystals. In an earlier[1] paper it was shown that the resonance condition in a uniaxial or cubic antiferromagnetic crystal at 0°K is given by

$$\omega/\gamma=H_0\pm[H_A(2H_E+H_A)]^{\frac{1}{2}}, \tag{1}$$

when the static magnetic field H_0 is parallel to the domain axis; here H_E is the Weiss exchange field, H_A is the anisotropy field, and $\gamma=ge/2mc$ is the magnetomechanical ratio; g is the spectroscopic splitting factor.

A number of the better-known antiferromagnetic crystals have Curie temperatures of the order of magnitude of 100°K, and we may for these estimate very roughly that $H_E\sim10^6$ oersteds and $H_A\sim10^3$ oersteds, leading to a resonance frequency $\omega\sim5$ cm^{-1} for zero static field; to obtain resonance at a microwave frequency ~1 cm^{-1} it is necessary to apply a static field $H_0\sim5\times10^4$ oersteds. We know also of antiferromagnetic crystals with lower Curie temperatures, and resonance experiments have been performed at Leiden[2] on a crystal ($CuCl_2\cdot2H_2O$) with a Curie temperature ~5°K, so that here we should expect a zero field splitting equivalent to 10^3 to 10^4 oersteds. It seems that a splitting of this nature was observed, although the published account of the measurements is very incomplete. Our frequency estimates may be quite badly off, as there are few data from which to estimate the anisotropy field.

In Sec. II of this paper we carry out a classical calculation of the resonance frequencies for a cubic antiferromagnetic crystal at 0°K with the static field making an arbitrary angle with the axis of the antiferromagnetic domains. We consider the nature of the normal modes of motion of the spin systems, and calculate rf susceptibilities, which may be very considerably larger than in paramagnetic salts.

In Sec. III the theory is extended to arbitrary temperatures, and demagnetizing effects are considered. As we approach the Curie temperature the phenomenon comes to resemble ordinary paramagnetic resonance. The demagnetizing effect of the shape of the specimen enters in a subtle way, but is only of interest when the exchange field is not too much greater than the demagnetizing field.

In Sec. IV we consider the effect of the anisotropy energy when there is an anisotropic coupling between the antiferromagnetic sublattices. We also extend the theory to crystals of orthorhombic symmetry, of some experimental interest. Section V is concerned with line widths, and it is pointed out that polycrystalline specimens are not well-suited to experimental work because of the large orientational broadening.

In Sec. VI the resonance Eq. (1) is derived quantum-mechanically, employing the powerful and general method used by Van Vleck[3] in treating ferromagnetic resonance. It is reassuring that the quantum-mechanical calculation agrees with the classical results and introduces no new features. We may note that both the classical and quantum calculations given in this paper utilize implicitly the two sublattice model of antiferromagnetics which was employed previously by Van Vleck[4] for static calculations. This model assumes in effect that the antiferromagnetic ground state is adequately approximated by a spin function of the character $\alpha_1\beta_2\alpha_3\beta_4\alpha_5\cdots$; it is satisfying in this connection that Anderson[5] has proved that the exact ground-state eigenvalue must be close to the energy of the two sublattice approximation. For structures such as MnO, where the paramagnetic ions form a face-centered cubic lattice, there are eight sublattices. Shull, Strauser, and Wollan[6] have found by neutron diffraction that four

[1] C. Kittel, Phys. Rev. **82**, 565 (1951).
[2] Poulis, van den Handel, Ubbink, Poulis, and Gorter, Phys. Rev. **82**, 552 (1951).
[3] J. H. Van Vleck, Phys. Rev. **78**, 266 (1950).
[4] J. H. Van Vleck, J. Chem. Phys. **9**, 85 (1941). For a general review of antiferromagnetism the reader is referred to J. H. Van Vleck, J. phys. et radium **12**, 262 (1951).
[5] P. W. Anderson, Phys. Rev. **83**, 1260 (1951).
[6] Shull, Strauser, and Wollan, Phys. Rev. **83**, 333 (1951).

sublattices are magnetized in one direction, and four in the opposite direction. In this case in a resonance experiment we deal only with the two groups of spins, as within each group every spin sees the same exchange and anisotropy fields. In Sec. VII we review the experimental situation.

II. CLASSICAL CALCULATION FOR 0°K, CUBIC SYMMETRY

A. Static Field Parallel to Preferred Axis

This case has been briefly treated in an earlier paper.[1] The equations of motion for the two sublattices are

$$\begin{aligned}
dM_1^x/dt &= \gamma(H^z+H_0+H_A-\lambda M_2^z)M_1^y \\
&\quad -\gamma(H^y-\lambda M_2^y)M_1^z, \\
dM_2^x/dt &= \gamma(H^z+H_0-H_A-\lambda M_1^z)M_2^y \\
&\quad -\gamma(H^y-\lambda M_1^y)M_2^z, \\
dM_1^y/dt &= -\gamma(H^z+H_0+H_A-\lambda M_2^z)M_1^x \\
&\quad +\gamma(H^x-\lambda M_2^x)M_1^z, \\
dM_2^y/dt &= -\gamma(H^z+H_0-H_A-\lambda M_1^z)M_2^x \\
&\quad +\gamma(H^x-\lambda M_1^x)M_2^z, \\
dM_1^z/dt &= \gamma(H^y-\lambda M_2^y)M_1^x-\gamma(H^x-\lambda M_2^x)M_1^y, \\
dM_2^z/dt &= \gamma(H^y-\lambda M_1^y)M_2^x-\gamma(H^x-\lambda M_1^x)M_2^y,
\end{aligned} \tag{2}$$

where H_0 is the static field, assumed in the preferred (z) direction, $\mathbf{H}$ is the rf field; $-\lambda\mathbf{M}_2$ and $-\lambda\mathbf{M}_1$ are the exchange fields acting on $\mathbf{M}_1$ and $\mathbf{M}_2$ respectively; H_A and $-H_A$ are the anisotropy fields; and $\gamma=ge/2mc$. For antiferromagnetism λ is positive. We assume small deflections from the preferred axis so that the anisotropy may be treated as an effective field directed along the axis.

To find the resonance frequency we set $\mathbf{H}=0$. Products like $M_2^yM_1^z$ can be taken as zero since the magnetization is assumed to be predominantly in the z direction.[7] Let $H_E=\lambda M_1^z=-\lambda M_2^z$ and assume M_1^x, M_2^x, M_1^y, M_2^y all proportional to $\exp(i\omega t)$.[8] The resonance frequencies are then the eigenvalues of the resonance matrix $\mathbf{A}$:

$$\mathbf{A}=\begin{vmatrix} 0 & 0 & -ia & -ib \\ 0 & 0 & ib & -id \\ ia & ib & 0 & 0 \\ -ib & id & 0 & 0 \end{vmatrix}, \tag{3}$$

where $a=\gamma(H_0+H_A+H_E)$, $b=\gamma H_E$, $d=\gamma(H_0-H_A-H_E)$. The matrix equation

$$\mathbf{AS}=\mathbf{SA}' \quad (\mathbf{A}' \text{ diagonal}) \tag{4}$$

is satisfied by

$$\mathbf{S}=\begin{vmatrix} 1 & 1 & 1 & 1 \\ -\eta^{-1} & -\eta^{-1} & -\eta & -\eta \\ i & -i & i & -i \\ -i\eta^{-1} & i\eta^{-1} & -i\eta & i\eta \end{vmatrix}, \tag{5}$$

where

$$\eta H_E=H_A+H_E+[H_A(2H_E+H_A)]^{\frac{1}{2}}, \tag{6}$$

and hence the columns of $\mathbf{S}$ are eigenvectors of $\mathbf{A}$ and represent the four normal modes N_i. The corresponding resonance frequencies (eigenvalues) are

$$\begin{aligned}
\omega_1 &= -\omega_2=\gamma H_0+\gamma[H_A(2H_E+H_A)]^{\frac{1}{2}}, \\
\omega_3 &= -\omega_4=\gamma H_0-\gamma[H_A(2H_E+H_A)]^{\frac{1}{2}}.
\end{aligned} \tag{7}$$

The normal modes are shown in Fig. 1. N_1 and N_2 are equivalent and represent $\mathbf{M}_1$ and $\mathbf{M}_2$ in a circular clockwise precession of frequency ω_1 about $+z$, i.e., looking along the z axis one sees $\mathbf{M}_1$ and $\mathbf{M}_2$ describing unequal size circles in the same sense.

To obtain a better picture of this motion consider Fig. 1(a) with $H_0=0$. If the precessions are to be in the same sense and of magnitude ω, and with $M=|\mathbf{M}_1|=|\mathbf{M}_2|$, the following equations of motion must be satisfied:

$$\begin{aligned}
\omega M\sin\theta_1 &= \gamma H_A M\sin\theta_1+\gamma M^2\lambda\sin(\theta_1-\theta_2), \\
\omega M\sin\theta_2 &= -\gamma H_A M\sin\theta_2+\gamma M^2\lambda\sin(\theta_1-\theta_2).
\end{aligned}$$

Assuming θ_1 and θ_2 small, these equations may be solved for ω and for θ_1/θ_2. They yield

$$\omega=\gamma[H_A(2H_E+H_A)]^{\frac{1}{2}}; \quad \theta_1/\theta_2=\eta.$$

When H_0 is added, the precessional frequency of both $\mathbf{M}_1$ and $\mathbf{M}_2$ is simply increased by γH_0. Figure 1(b) can be interpreted in an analogous manner.

For $H_0=0$ we have degeneracy, and $\omega_1=\omega_4$, $\omega_2=\omega_3$. We may combine N_1 with N_4, N_2 with N_3 into the linear combinations P_i as follows:

$$\begin{aligned}
P_1 &= \eta N_1-N_4, \quad P_3=\eta N_1+N_4, \\
P_2 &= \eta N_2-N_3, \quad P_4=\eta N_2+N_3,
\end{aligned} \tag{8}$$

giving rise to a transformation

$$\mathbf{S}'=\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ i\epsilon & -i\epsilon & i\epsilon^{-1} & -i\epsilon^{-1} \\ -i\epsilon & i\epsilon & i\epsilon^{-1} & -i\epsilon^{-1} \end{vmatrix}, \tag{9}$$

where

$$\epsilon=(\eta+1)/(\eta-1)=+[(2H_E+H_A)/H_A]^{\frac{1}{2}}.$$

P_1 and P_2 are now equivalent, as are P_3 and P_4. In these modes $\mathbf{M}_1$ precesses clockwise, $\mathbf{M}_2$ counterclockwise about $+z$, i.e., looking along z one sees $\mathbf{M}_1$ and $\mathbf{M}_2$ describing equal size ellipses in opposite directions, with the ratio of major to minor axis equal to ϵ. A discussion of this motion is illuminating. In mode P_1 (or P_2) the vectors $\mathbf{M}_1$ and $\mathbf{M}_2$ are colinear as they cross the yz plane and each is acted upon by only its

[7] Also because, for quantum-mechanical reasons, we seek transitions for which M_1^z and M_2^z are constants of the motion—other transitions giving us multiples of the fundamental frequency. For a discussion of this point see J. H. Van Vleck, Phys. Rev. 74, 1168 (1948).

[8] They are actually proportional to $\sin\omega t$ or $\cos\omega t$. This can be taken care of, and the phase relations identified, by noting the complex amplitudes of the normal modes in, e.g., Eq. (5).

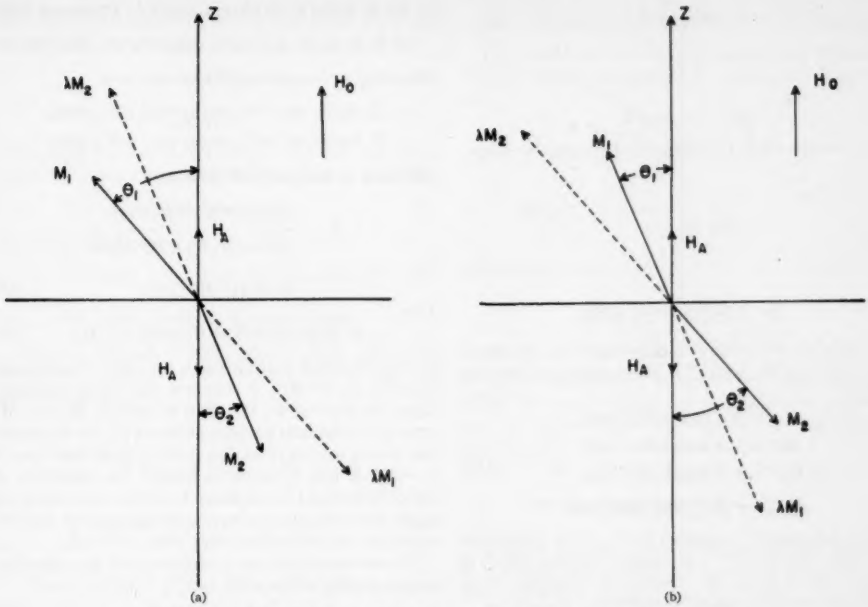


FIG. 1. (a) The normal modes N_1 and N_2. $\theta_1/\theta_2=\eta$. (b) The normal modes N_3 and N_4. $\theta_2/\theta_1=\eta$.

anisotropy field, H_A. One-quarter cycle later as the vectors cross the xz plane they will be non-collinear, pointing in the same direction along x, and hence making an angle with H_E of twice that made with H_A. Each will be acted on by a field equal to $2H_E+H_A$. We compare the situation to ferromagnetic resonance in a plane surface, where the fields seen one-quarter cycle apart are B and H, and the resonance frequency is $\gamma(BH)^{\frac{1}{2}}$. Similarly we have here $\pm\gamma[H_A(2H_E+H_A)]^{\frac{1}{2}}$. The $\pm$ arises here since we deal with two magnetizations precessing in opposite directions. In mode P_3 (or P_4) we have the same argument, except the vectors are colinear as they cross the xz plane. One can, of course, form analogous combinations of the N_i using arbitrary axes x, y, resulting in a similar motion. That is, the elliptical paths cut out in the xy plane are of arbitrary orientation.

B. Susceptibilities

By solving Eq. (2) with the rf field $\mathbf{H}$ left in, it is found that, for a circularly polarized field $H^-=H^x-iH^y$, and defining $M^-=M^x-iM^y$, $M=|\mathbf{M}_1|=|\mathbf{M}_2|$, the susceptibility is

$$\chi^-=M^-/H^-=-2\gamma^2H_AM/[(\omega-\omega_1)(\omega-\omega_3)], \quad (10)$$

with ω_1 and ω_3 as given by Eq. (7). For a linearly polarized rf field

$$\chi^x=\frac{M^x}{H^x}=\frac{-2\gamma^2H_AM(\omega^2+\omega_1\omega_3)}{(\omega^2-\omega_1^2)(\omega^2-\omega_3^2)}. \quad (11)$$

These susceptibilities may be quite large compared to those of paramagnetic resonance. Assuming a relaxation frequency Δ we have for the antiferromagnetic susceptibility at $\omega=\omega_1$:

$$\chi_{AF}^-\cong-i\gamma H_AM/[(2H_AH_E)^{\frac{1}{2}}\Delta],$$

and for the paramagnetic susceptibility at $\omega=\gamma H_0$:

$$\chi_P^-\cong-i\gamma M_P/\Delta.$$

If the Δ's are the same, and with $M_P\approx\lambda^{-1}H_0$ (near the Curie point the static susceptibility is λ^{-1}), we have for the ratio of the rf susceptibilities at 0°K and just above the Curie point:

$$\chi_{AF}^-/\chi_P^-\cong(H_AH_E)^{\frac{1}{2}}/H_0; \quad (12)$$

this ratio is ~10 for $H_0\sim10^4$, $H_E\sim10^6$, $H_A\sim10^4$ oersteds.

C. Static Field Perpendicular to Preferred Axis

Here $\mathbf{M}_1$ and $\mathbf{M}_2$ find new equilibrium positions at angles ϕ from the preferred axes. Assuming now H_0 along the x direction we introduce two new sets of

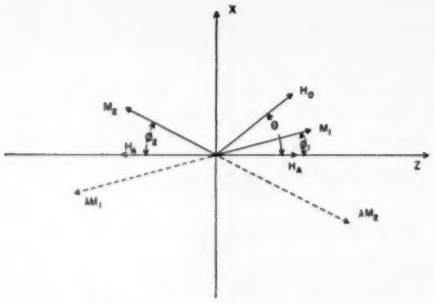


FIG. 2. H_0 at arbitrary angle.

axes: $x'y'z'$ for $\mathbf{M}_1$ (with z' in the equilibrium direction) and $x''y''z''$ for $\mathbf{M}_2$ (with z'' in the equilibrium direction of $\mathbf{M}_2$):

$$\begin{aligned} M_1{}^{x'} &= M_1{}^x \cos\phi - M_1{}^y \sin\phi, \\ M_1{}^{y'} &= M_1{}^x \sin\phi + M_1{}^y \cos\phi, \\ M_2{}^{x''} &= M_2{}^x \cos\phi + M_2{}^y \sin\phi, \\ M_2{}^{y''} &= -M_2{}^x \sin\phi + M_2{}^y \cos\phi. \end{aligned} \tag{13}$$

This transformation is applied to Eq. (2) in absence of $\mathbf{H}$. Terms like $M_2{}^y M_1{}^{z'}$ are dropped. Setting $dM_1{}^{x'}/dt = dM_2{}^{y''}/dt = 0$ gives the equilibrium condition on ϕ and hence the fundamental frequencies.[7] This gives

$$-H_0 \cos\phi + H_A \sin\phi + 2H_E \cos\phi \sin\phi = 0.$$

When $H_0 \ll H_E$, and since $H_A \ll H_E$, ϕ will be very small and hence

$$\phi \cong H_0/2H_E, \tag{14}$$

which gives for the static susceptibility

$$\chi_\perp = (M_{1\perp} + M_{2\perp})/H_0 = \lambda^{-1}, \tag{15}$$

a well-known result.

On using Eq. (14) the transformed equations of motion simplify to equations similar to Eq. (2), but in different coordinate systems. The resulting resonance matrix yields the frequencies

$$(\omega/\gamma)^2 = 2H_E H_A + H_A{}^2 + \tfrac{1}{2}H_0{}^2 \pm (H_E + H_A)(H_0{}^2/2H_E). \tag{16}$$

Recalling that since ϕ must be small for our approximation, H_0 must be $\ll H_E$, and using $H_A \ll H_E$:

$$\omega/\gamma \cong \pm[2H_E H_A + H_0{}^2]^{\frac{1}{2}}, \tag{17a}$$

and

$$\omega/\gamma \cong \pm(2H_E H_A)^{\frac{1}{2}}. \tag{17b}$$

On solving the equations of motion for the rf susceptibilities one finds that, if the rf field is in the usual y direction, only the frequencies (17a) will be excited. To excite the frequencies (17b) the rf field must be in the direction of H_0.[8a]

[8a] We are indebted to Professor T. Nagamiya of Osaka University for calling our attention to the existence of the frequencies (17b).

D. Static Field at Arbitrary Angle to Preferred Axis

Let H_0 be in the xz plane at angle θ to $+z$. (See Fig. 2.)

Balancing torques at equilibrium we have

$$\begin{aligned} H_0 \sin(\theta - \phi_1) &= H_E \sin(\phi_1 + \phi_2) + H_A \sin\phi_1, \\ H_0 \sin(\theta + \phi_2) &= H_E \sin(\phi_1 + \phi_2) + H_A \sin\phi_2. \end{aligned}$$

Assuming ϕ_1 and ϕ_2 small these give

$$\phi_{1,2} = \frac{H_0 \sin\theta(H_A \mp H_0 \cos\theta)}{H_A{}^2 + 2H_A H_E - H_0{}^2 \cos^2\theta}.$$

If

$$H_0{}^2 \ll 2H_A H_E/\cos^2\theta, \tag{18}$$

Then

$$\phi_{1,2} \cong H_0 \sin\theta(H_A \mp H_0 \cos\theta)/2H_A H_E. \tag{19}$$

It is to be noted that although Eq. (18) only requires $H_0 \ll H_E$ at $\theta \sim 90°$, it requires the more stringent $H_0 \ll (2H_A H_E)^{\frac{1}{2}}$ at $\theta < 45°$. This is because $\mathbf{M}_1$ and $\mathbf{M}_2$ attempt to become perpendicular to H_0 for minimum free energy and have weaker fields to overcome when θ is $\sim 45°$ or less. In order to remove the restriction of Eq. (18) it would be necessary to depart from the small angle approximation we have used throughout, and the equations would become much more involved.

The resonance frequency may be found by a transformation similar to Eq. (13). Let

$$y = H_0{}^2 \sin^2\theta / 2H_A H_E. \tag{20}$$

Thus the resonance frequencies are

$$(\omega/\gamma^2)^2 \cong 2H_E H_A + H_0{}^2[\cos^2\theta(1-y)^2 + \tfrac{1}{2}\sin^2\theta] \pm \tfrac{1}{2}H_0[H_0{}^2 \sin^4\theta + 16H_E H_A \cos^2\theta(2+y)(1-y)^2]^{\frac{1}{2}}. \tag{21}$$

The normal modes and rf susceptibilities are quite involved. With the rf field in the usual y direction all frequencies should be observed if θ is appreciably less than 90°.

III. EXTENSION TO ARBITRARY TEMPERATURES, CUBIC SYMMETRY

A. Static Field Parallel to Preferred Axis

We assume the ordering sufficiently strong for vectors $\mathbf{M}_1$ and $\mathbf{M}_2$ to have a meaning. First the case of field along the preferred axis is worked out, then the methods of Sec. II are used to extend the result to arbitrary field direction. The equations of motion (2) are still valid providing we can assume H_A the same for both sublattices. This is adequate since near the Curie point where, as we can see from microscopic models, the anisotropy fields will be noticeably different, they will also be negligibly small. If H_0 is large enough to cause an appreciable difference in the two anisotropy fields, it will also be large enough to dominate the resonance. We assume $H_A/H_E \ll 1$ at all temperatures.

The analog of Eq. (7) when $|M_1{}^{z'}| \neq |M_2{}^{z'}|$ is

$$\omega/\gamma = H_0 - \tfrac{1}{2}\lambda(M_1{}^{z'} + M_2{}^{z'}) \pm \{H_A[H_A + \lambda(M_1{}^{z'} - M_2{}^{z'})] + (\lambda/2)^2(M_1{}^{z'} + M_2{}^{z'})^2\}^{\frac{1}{2}}. \tag{22}$$

This may be simplified by noting that since $M_1^x+M_2^x = \chi_{11}H_0$, and on using Eq. (15),

$$\lambda(M_1^x+M_2^x)=\alpha H_0, \tag{23}$$

where

$$\alpha=\chi_{11}/\chi_\perp; \tag{24}$$

α increases from 0 to 1 between 0°K and the Curie point. We define a temperature-dependent average H_E:

$$2H_E=\lambda(M_1^z-M_2^z). \tag{25}$$

Thus Eq. (22) becomes

$$\omega/\gamma\cong H_0(1-\alpha/2)\pm[2H_EH_A+(\alpha/2)^2H_0^2]^{\frac{1}{2}}. \tag{26}$$

B. Static Field at Arbitrary Angle

Figure 2 applies, except $|\mathbf{M}_1|>|\mathbf{M}_2|$. Balancing torques at equilibrium and solving for ϕ_1 and ϕ_2, assuming them small

$$\phi_{1,2}=\frac{H_0\sin\theta[H_A\mp(1-\alpha)H_0\cos\theta]}{2H_AH_E+H_A^2-(1-\alpha)H_0^2\cos^2\theta}.$$

If

$$H_0^2\ll 2H_AH_E/[(1-\alpha)\cos^2\theta], \tag{27}$$

then

$$\phi_{1,2}\cong H_0\sin\theta[H_A\mp(1-\alpha)H_0\cos\theta]/2H_AH_E. \tag{28}$$

Equations (27) and (28) are analogous to (18) and (19) and the remarks given below the latter apply here also.[8b] Proceeding as in the 0°K case we find the resonance frequencies by solving the characteristic equation. The result is quite complicated algebraically unless we make the assumptions

$$H_A(T)\ll H_E(T) \tag{29}$$

and

$$H_0^2\ll ZH_A(T)H_E(T). \tag{30}$$

Under these restrictions the resonance frequencies are

$$\begin{aligned}(\omega/\gamma)^2\cong{}&2H_AH_E+\tfrac{1}{2}H_0^2[1+(1+\alpha)^2\cos^2\theta]\\&\pm\tfrac{1}{2}H_0\{8H_AH_E(Z-\alpha)^2\cos^2\theta\\&+H_0^2[\sin^4\theta+\cos^2\theta(Z-\alpha)^2(Z\sin^2\theta+\alpha^2\cos^2\theta)]\}^{\frac{1}{2}}.\end{aligned} \tag{31}$$

This equation, in the form of H_0 as a function of ω, has been given by Nagamiya.[9] It is to be noted that his α is equal to our $(1-\alpha)$.

[8b] It should be noted that $2H_AH_E/(1-\alpha)$ may well *increase* with increasing temperature. If it does, it should take larger fields H_0 as the temperature rises to cause the parallel susceptibility suddenly to jump to a high value, indicating that the magnetizations have flopped over to the hard axis (or to a perpendicular axis in case of cubic symmetry). Such an increase with temperature has been found at Leiden in $CuCl_2\cdot 2H_2O$. See reference 2. This type of experiment, incidentally, offers an independent method of determining some of the quantities entering into the resonance equations, since from free energy considerations it is seen that the threshhold strength of field applied parallel to the easy axis and causing a flop is

$$H_0^2=2H_AH_E/(1-\alpha).$$

[9] T. Nagamiya, Prog. Theor. Phys. **6**, 342 (1951).

C. Temperature Dependence of $\alpha(=\chi_{11}/\chi_\perp)$

Van Vleck[4] has derived expressions for the susceptibility of antiferromagnetics in small fields. His results are

$$\alpha/2=(1+A)^{-1}, \tag{32}$$

where, in particular, for spin $S=1/2$

$$A=\frac{T/T_c}{1-(M/M_0)^2}$$

and for spin $S\to\infty$

$$A=\frac{T/3T_c}{1-(M/M_0)^2-2T/3T_c}.$$

In these relations T_c is the Curie temperature and M/M_0 is the ratio of the spontaneous inner magnetization of one sublattice at temperature T to that at absolute zero. This ratio rises rapidly to 1 as T goes below T_c and is plotted in books on ferromagnetism.

D. Demagnetizing Effects

We show that the central maximum of the resonance line is essentially unaffected except in materials with very low Curie points (small exchange fields). Effects on line width are considered in Sec. V.

Assuming an elliptical specimen with demagnetizing coefficients N^x, N^y, N^z, we take account of demagnetizing effects in the resonance equations by replacing $\mathbf{H}$ in Eq. (2) by

$$\begin{aligned}\mathbf{H}-[&\mathbf{i}N^x(M_1^x+M_2^x)+\mathbf{j}N^y(M_1^y+M_2^y)\\&+\mathbf{k}N^z(M_1^z+M_2^z)].\end{aligned}$$

The general solution at arbitrary temperature and angles of H_0 is quite involved algebraically. We give two special cases.

(*1*) *0°K, H_0 along preferred axis.*—

$$\begin{aligned}(\omega/\gamma)^2=H_0^2+H_A^2+H_AH_E[2+(N^y+N^z)/\lambda]\\\pm\{4H_0^2H_A^2+4H_0^2H_AH_E[2+(N^y+N^z)/\lambda]\\+H_A^2H_E^2(N^y-N^z)/\lambda\}^{\frac{1}{2}}.\end{aligned} \tag{33}$$

This reduces, for $H_0=0$, to

$$\omega/\gamma=\pm[2H_AH_E(1+N^z\lambda^{-1})]^{\frac{1}{2}} \tag{34a}$$

and

$$\pm[2H_AH_E(1+N^y\lambda^{-1})]^{\frac{1}{2}}. \tag{34b}$$

With $N^y\neq N^z$ there are four resonance frequencies. The difference between Eq. (34a) and Eq. (34b) may not be resolveable, but may serve only to broaden the lines by $\Delta\omega/\omega\sim(N^z-N^y)/2\lambda$. In MnO, e.g., where $\lambda^{-1}=\chi_\perp\sim 10^{-4}$ we have $\Delta\omega/\omega\sim 10^{-3}$ (or smaller). If the Curie point is 1°K, then $\Delta\omega/\omega\sim 10^{-1}$. The ideal shape for single-crystal experiments is one with $N^x=N^y$, i.e., with cylindrical symmetry. The normal modes of Eq. (34) are the same as those given by Eq. (42) of Sec. IV.

(2) $N^x=N^y=N$, *arbitrary temperature, and arbitrary direction of* H_0.—We have $N^z=4\pi-2N$. By means of the substitutions[10]

$$H_0''=H_0[1+(3N-4\pi)\alpha/\lambda],$$
$$2H_E''=(\lambda+N)(M_1^z-M_2^z),$$

we restore the equations of motion (2) in absence of demagnetizing factors, but now in terms of the primed fields. Hence Eq. (31) now holds with H_0 replaced by H_0'' and H_E replaced by H_E''. The change is not important if $\lambda\gg N$. If $N=4\pi/3$ (spherical sample) the only change is in the size of effective H_E.

IV. MACROSCOPIC DISCUSSION OF ANISOTROPY

A. Cubic Symmetry

Since the anisotropy is of extreme importance in determining antiferromagnetic resonance, some of the possibilities are examined from a macroscopic point of view in this section. Microscopic analogs, from a quadrupolar-coupling point of view, are given in Sec. VI.

For cubic symmetry we define macroscopic anisotropy constants in the usual way. Only quartic (or higher) forms in the direction cosines of the magnetization are consistent with cubic symmetry. However, for small displacements from a preferred axis, these forms will have quadratic terms which give rise to effective anisotropy fields.

The form

$$f_A=(K_1/2)(\alpha_1^2\beta_1^2+\alpha_1^2\gamma_1^2+\beta_1^2\gamma_1^2+\alpha_2^2\beta_2^2+\alpha_2^2\gamma_2^2+\beta_2^2\gamma_2^2),$$

where α_1, β_1, γ_1 and α_2, β_2, γ_2 are the direction cosines of $\mathbf{M}_1$ and $\mathbf{M}_2$ respectively, has a quadratic term for small displacements from the z-axis

$$f_A'\cong(K_1/2)(\alpha_1^2+\beta_1^2+\alpha_2^2+\beta_2^2). \tag{35a}$$

This is analogous to ferromagnetics. In addition, we have in antiferromagnetics the presence of forms depending upon the relative positions of $\mathbf{M}_1$ and $\mathbf{M}_2$, involving terms coupling the direction cosines. For small displacements from the z-axis these can all be reduced to

$$f_A''\cong(K_2/2)(\alpha_1^2+\beta_1^2+\alpha_2^2+\beta_2^2) \tag{35b}$$

and

$$f_A'''\cong K_3(\alpha_1\alpha_2+\beta_1\beta_2). \tag{35c}$$

We may readily identify the H_A used in Secs. II and III:

$$H_AM_1^x=(\partial/\partial\alpha_1)(f_A'+f_A'')=(K_1+K_2)M_1^x/M;$$

and we now have an additional field h_A

$$h_AM_1^x=\partial f_A'''/\partial\alpha_1=K_3M_2^x/M,$$

which adds new terms to the equations of motion (2).

The original equations of motion are restored by redefining

$$H_E'''=H_E+(K_3/M),\quad H_A'''=(K_1+K_2-K_3)/M. \tag{36}$$

The change in H_E may be neglected. Thus all the results of Secs. II and III hold with H_AH_E replaced by $\lambda(K_1+K_2-K_3)$. We have assumed anisotropy fields of the same magnitude acting on the two sublattices, and we have made use of Eq. (25).

As redefined H_A''' cannot become negative as can be seen from the following argument. For small displacements from the z-axis, and with $H_0=0$, the free energy is

$$F\cong\lambda M_1M_2[\alpha_1\alpha_2+\beta_1\beta_2-(1-\tfrac{1}{2}\alpha_1^2-\tfrac{1}{2}\beta_1^2)(1-\tfrac{1}{2}\alpha_2^2-\tfrac{1}{2}\beta_2^2)]+[(K_1+K_2)/2](\alpha_1^2+\alpha_2^2+\beta_1^2+\beta_2^2)+K_3(\alpha_1\alpha_2+\beta_1\beta_2).$$

This may be written as

$$F\cong(\lambda M_1M_2+K_3)[\alpha_1\alpha_2+\beta_1\beta_2-(1-\tfrac{1}{2}\alpha_1^2-\tfrac{1}{2}\beta_1^2)(1-\tfrac{1}{2}\alpha_2^2-\tfrac{1}{2}\beta_2^2)]+[(K_1+K_2-K_3)/2](\alpha_1^2+\alpha_2^2+\beta_1^2+\beta_2^2),$$

which immediately justifies Eq. (36). In order that F have a minimum at $\alpha_1=\alpha_2=\beta_1=\beta_2=0$, as postulated when we assume z a preferred axis, F must be a "definite" quadratic form. The condition for this is

$$\frac{\partial^2F}{\partial\alpha_1^2}\cdot\frac{\partial^2F}{\partial\alpha_2^2}>\left(\frac{\partial^2F}{\partial\alpha_1\partial\alpha_2}\right)^2$$

or

$$(\lambda M_1M_2+K_1+K_2)^2>(\lambda M_1M_2+K_3)^2,$$

so that

$$K_1+K_2>K_3. \tag{37}$$

Quantum-mechanical expressions for K_1, K_2, and K_3 are given, in terms of quadrupolar coupling constants, in Sec. VI.

B. Temperature Dependence of $H_E'''H_A'''$ $(\cong K_1+K_2-K_3)$

Here we have little to go by. As noted by Van Vleck[11] his calculations for ferromagnetic anisotropy exhibit the general trend with temperature that is observed experimentally, but do not give anywhere near as rapid a diminishment of anisotropy near the Curie point as is observed. In addition, some ferromagnetics, e.g., MnBi, exhibit an *increase* of anisotropy over a wide range of increasing temperature,[12] quite contrary to the predictions of the quadrupolar coupling model. Presumably antiferromagnetics are subject to the same vagaries.

[10] These substitutions can also be found by considering the free energy (less anisotropy):

$$F=-H_0(M_1^x+M_2^x)+\lambda\mathbf{M}_1\cdot\mathbf{M}_2+\tfrac{1}{2}N(M_1^x+M_2^x)^2+\tfrac{1}{2}N(M_1^y+M_2^y)^2+\tfrac{1}{2}(4\pi-2N)(M_1^z+M_2^z)^2;$$

which may be rewritten

$$F=\tfrac{1}{2}N(\mathbf{M}_1^2+\mathbf{M}_2^2)-[H_0+\tfrac{1}{2}(3N-4\pi)(M_1^z+M_2^z)](M_1^z+M_2^z)+(\lambda+N)\mathbf{M}_1\cdot\mathbf{M}_2.$$

[11] J. H. Van Vleck, Phys. Rev. **52**, 1178 (1937).

[12] C. Guillaud, *Ferromagnétisme des alliages binaires de manganèse*, thesis, Strasbourg, 1943.

Assuming, however, that the quadrupolar model is followed, we have from Sec. VI below that the temperature dependent part of $K_1+K_2-K_3$ goes approximately as

$$[S(S+1)-3\langle(S_{1j'})^2\rangle_{Av}][\langle(S_{1j''})_{Av}]^2, \tag{38}$$

where

$$\langle S_{1j'}\rangle_{Av}=B_1(\theta)=\sum_{m=-S}^{S} me^{m\theta}\Big/\sum_{m=-S}^{S} e^{m\theta},$$

$$\langle(S_{1j'})^2\rangle_{Av}=B_2(\theta)=\sum_{m=-S}^{S} m^2e^{m\theta}\Big/\sum_{m=-S}^{S} e^{m\theta},$$

with

$$\theta=g\beta H_0/kT.$$

$B_1(\theta)$ is the Brillouin function, and it is to be noted that

$$B_2(\theta)=[B_1(\theta)]^2+dB_1/d\theta.$$

Further details on this model as applied to ferromagnetics may be found in Van Vleck's papers.[8,11]

C. Noncubic, but Orthorhombic Symmetry

This case includes $CuCl_2\cdot 2H_2O$ which has been investigated at Leiden.[2] The anisotropy may be a quadratic function of the direction cosines without the small angle approximation. We have as the most general quadratic term, after subtracting out the rotational invariant dot product of the magnetization vectors (which acts as an exchange term)

$$f_A=(K_1'/2)(\alpha_1^2+\alpha_2^2)+(K_2'/2)(\beta_1^2+\beta_2^2)+K_3'\alpha_1\alpha_2+K_4'\beta_1\beta_2. \tag{39}$$

We choose the z axis as the preferred (direction of easiest magnetization). Hence, using arguments similar to those leading to Eq. (37) we must have

$$K_1'>K_3' \quad \text{and} \quad K_2'>K_4'. \tag{40}$$

We treat two special cases, and include demagnetizing factors.

(1) 0°K, $H_0=0$

The resonance matrix corresponding to Eq. (3) is

$$\mathbf{A}_3=\begin{vmatrix} 0 & 0 & -ia & -ib \\ 0 & 0 & ib & ia \\ ic & id & 0 & 0 \\ -id & -ic & 0 & 0 \end{vmatrix}, \tag{41}$$

where

$$a=\gamma H_E(1+N^y\lambda^{-1})+\gamma K_2'/M,$$
$$b=\gamma H_E(1+N^y\lambda^{-1})+\gamma K_4'/M,$$
$$c=\gamma H_E(1+N^x\lambda^{-1})+\gamma K_1'/M,$$
$$d=\gamma H_E(1+N^x\lambda^{-1})+\gamma K_2'/M.$$

The diagonalizing matrix becomes

$$\mathbf{S}_3=\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ i\mu & -i\mu & i\nu^{-1} & -i\nu^{-1} \\ -i\mu & i\mu & i\nu^{-1} & -i\nu^{-1} \end{vmatrix}, \tag{42}$$

where

$$\mu=+[(c+d)/(a-b)]^{\frac{1}{2}}, \quad \nu=+[(a+b)/(c-d)]^{\frac{1}{2}}.$$

This is closely analogous to the $\mathbf{S}'$ of Eq. (9). The eigenvalues are

$$\omega_1=-\omega_2=+[(a-b)(c+d)]^{\frac{1}{2}},$$
$$\omega_3=-\omega_4=+[(c-d)(a+b)]^{\frac{1}{2}}. \tag{43}$$

The normal modes P_i are similar to those of Eq. (9) except that the elliptical paths in the xy plane are no longer of arbitrary orientation, and are of different frequencies according to whether the major axes are in the x or in the y directions.

The susceptibilities are interesting. For linearly polarized rf fields

$$\chi^x=-2\gamma M(a-b)/(\omega^2-\omega_1^2),$$
$$\chi^y=-2\gamma M(c-d)/(\omega^2-\omega_3^2).$$

Thus we have the somewhat unusual result that only the resonance frequency ω_1 should be observed with the rf field in the x direction, and only ω_3 with the rf in the y direction. Also, assuming the ω_1 and ω_3 lines to be of similar shape, the ratio of amplitudes should go as $(a-b)/(c-d)$, or as the ratio of the effective anisotropy fields for x and y displacements.

(2) 0°K, H_0 along preferred axis.

The resonance frequencies are

$$\omega^2=\gamma^2H_0^2+ac-bd\pm[(a+c-b-d)(a+b+c+d)H_0^2+(bc-ad)^2]^{\frac{1}{2}}. \tag{44}$$

The normal modes are too complicated to admit of a simple physical interpretation. All frequencies should be observed in both the x and y directions, although with quite disparate intensities as long as H_0 is small.

V. LINE WIDTHS

Powder samples should give according to Eq. (31) a fractional line width $\Delta H_0/H_0\sim 1$. With H_0 fixed and ω varied the broadening will be $\Delta\omega/\omega\sim H_0/(2H_EH_A)^{\frac{1}{2}}$. Since the microcrystals have random demagnetizing factors, there is according to Eq. (34) an additional contribution to the width of $\Delta\omega/\omega\sim 10\lambda^{-1}=10\chi_\perp$. This would vary from sample to sample, being $\sim 10^{-3}$ in MnO and becoming quite large for samples with low Curie points.

Single crystals cut cylindrically, as noted in Sec. III D, are the ideal specimens for studying antiferromagnetic resonance. The crystals should be carefully annealed in order to minimize lattice dislocations. A high density of dislocations may effectively smear out the antiferromagnetic crystal pattern, introducing many extra sets of spins in one or the other direction, and causing widely varying anisotropy and exchange forces. This problem is in some ways peculiar to antiferromagnetics, for in ferromagnetics (a) the spin pattern is not essentially changed by dislocations, and (b) the resonance frequency is not a critical function of the product of exchange and anisotropy fields.

In both powders and crystals there will also be a broadening by various relaxation mechanisms. The dipolar broadening should be small because of the Van Vleck-Gorter process of exchange narrowing[13] which is independent of sign of exchange integral and which is important when the exchange is strong.

VI. QUANTUM-MECHANICAL CALCULATION, CUBIC SYMMETRY

We follow the method used by Van Vleck[3] in treating ferromagnetic resonance. Only the case H_0 along the preferred axis will be treated; the general case follows by using transformations as in Eq. (13). As in the classical treatment, a two sublattice picture is used. We neglect dipole-dipole terms as being relatively unimportant. The anisotropy is treated as a quadrupole-quadrupole coupling. As Van Vleck points out, this is merely an artifice, representing in a simple way the complex interaction among spins, orbits, and lattice. It leaves us with unknown quadrupolar coupling constants, and the best we can do is to identify these with macroscopic counterparts.

The Hamiltonian of the system is the sum of Zeeman, exchange, and quadrupolar terms

$$\mathcal{H}=\mathcal{H}_Z+\mathcal{H}_E+\mathcal{H}_Q. \tag{45}$$

We assume spins $S_{1j}{}^z$ of sublattice 1 directed opposite to H_0, so that $M_1{}^z$ is in the direction of H_0. Introducing exchange coupling constants A_{jk} and quadrupolar coupling constants K_{jk} we have

$$\mathcal{H}_Z=H_0g\beta(\textstyle\sum_j S_{1j}{}^z+\sum_j S_{2j}{}^z);$$

$$\mathcal{H}_E=\textstyle\sum_{k>j} A_{jk}\mathbf{S}_{1j}\cdot\mathbf{S}_{1k}+\sum_{k>j} A_{jk}'\mathbf{S}_{2j}\cdot\mathbf{S}_{2k}+\sum_{k,j} A_{jk}''\mathbf{S}_{1j}\cdot\mathbf{S}_{2k};$$

$$\mathcal{H}_Q=\textstyle\sum_{k>j} K_{jk}r_{jk}{}^{-4}(\mathbf{S}_{1j}\cdot\mathbf{r}_{jk})^2(\mathbf{S}_{1k}\cdot\mathbf{r}_{jk})^2+\sum_{k>j} K_{jk}'r_{jk}{}^{-4}(\mathbf{S}_{2j}\cdot\mathbf{r}_{jk})^2(\mathbf{S}_{2k}\cdot\mathbf{r}_{jk})^2+\sum_{k,j} K_{jk}''r_{jk}{}^{-4}(\mathbf{S}_{2j}\cdot\mathbf{r}_{jk})^2(\mathbf{S}_{1k}\cdot\mathbf{r}_{jk})^2.$$

Using the equations of motion and commutation relations

$$dS_{1j}{}^x/dt=(i/\hbar)[\mathcal{H},S_{1j}{}^x];\quad \text{etc.};$$

$$[S_{1j}{}^x,S_{1k}{}^y]=iS_{1j}{}^z\delta_{jk};\quad \text{etc.};$$

we obtain

$$\begin{aligned}&dS_1{}^z/dt=\textstyle\sum_j dS_{1j}{}^z/dt=Z+E+Q;\quad Z=-\gamma H_0S_1{}^y;\\ &\hbar E=\textstyle\sum_{k,j} A_{jk}''(S_{1j}{}^zS_{2k}{}^y-S_{1j}{}^yS_{2k}{}^z);\\ &\hbar Q=\textstyle\sum_{k\neq j}K_{jk}(\alpha_{jk}S_{1j}{}^x+\beta_{jk}S_{1j}{}^y+\gamma_{jk}S_{1j}{}^z)^2W_{jk}+\sum_{k,j} K_{jk}''(\alpha_{jk}S_{2j}{}^x+\beta_{jk}S_{2j}{}^y+\gamma_{jk}S_{2j}{}^z)^2W_{jk};\end{aligned} \tag{46}$$

where

$$W_{jk}=(\alpha_{jk}S_{1k}{}^x+\beta_{jk}S_{1k}{}^y+\gamma_{jk}S_{1k}{}^z)(\beta_{jk}S_{1k}{}^z-\gamma_{jk}S_{1k}{}^y)+(\beta_{jk}S_{1k}{}^z-\gamma_{jk}S_{1k}{}^y)(\alpha_{jk}S_{1k}{}^x+\beta_{jk}S_{1k}{}^y+\gamma_{jk}S_{1k}{}^z).$$

Here α_{jk}, β_{jk}, γ_{jk} are direction cosines of $\mathbf{r}_{jk}$ relative to the field axes, since the axis of quantization must be along H_0. When H_0 is along a preferred axis we may make the field axes coincide with the crystal axes, which is a great simplification. Similar equations of motion may be obtained for $S_2{}^x$, etc.

We make the approximations that below the Curie point we may replace $S_{1j}{}^z$ and $S_{2k}{}^z$ by their expectation values:

$$\langle S_{1j}{}^z\rangle_{\text{Av}}=-2M_1{}^z/Ng\beta;\quad \langle S_{2k}{}^z\rangle_{\text{Av}}=-2M_2{}^z/Ng\beta; \tag{47}$$

where $N/2$ is the number of atoms in a sublattice.

Assuming that local exchange fields can be replaced by average fields taken independently over successive clusters of nearest neighbors, next nearest neighbors, etc., we may define an average exchange parameter λ

$$\lambda=(2/Ng\beta\hbar\gamma)\textstyle\sum_j A_{jk}'' \tag{48}$$

independent of k. Then we have

$$E=\gamma\lambda M_2{}^zS_1{}^y-\gamma\lambda M_1{}^zS_2{}^y. \tag{49}$$

Note that this is independent of lattice structure.

To simplify Q we replace squares by average values for each atom, assumed independent. Thus we use the relations

$$\begin{aligned}&\langle(S_{1j}{}^x)^2\rangle_{\text{Av}}=\langle(S_{1j}{}^y)^2\rangle_{\text{Av}}=\tfrac{1}{2}[S(S+1)-\langle(S_{1j}{}^z)^2\rangle_{\text{Av}}];\\ &S_{1j}{}^xS_{1j}{}^y+S_{1j}{}^yS_{1j}{}^x=0;\\ &\tfrac{1}{2}(S_{1j}{}^xS_{1j}{}^z+S_{1j}{}^zS_{1j}{}^x)=[\langle S_{1j}{}^z\rangle_{\text{Av}}+\tfrac{1}{2}]S_{1j}{}^x;\end{aligned}$$

with similar relations for S_2. Also we take advantage of cubic symmetry to drop terms like[14]

$$\textstyle\sum_k K_{jk}(\alpha_{jk}{}^2-\beta_{jk}{}^2),\quad \sum_k K_{jk}\alpha_{jk}\beta_{jk},\quad \sum_k K_{jk}\alpha_{jk}{}^3\beta_{jk},\quad \text{etc.}$$

This eventually reduces Q to

$$\begin{aligned}\hbar Q=&[S(S+1)-3\langle(S_{1j}{}^z)^2\rangle_{\text{Av}}][\langle S_{1j}{}^z\rangle_{\text{Av}}+\tfrac{1}{2}]\\ &\times\textstyle\sum_k K_{jk}(\gamma_{jk}{}^4-3\beta_{jk}{}^2\gamma_{jk}{}^2)S_1{}^y\\ &+[S(S+1)-3\langle(S_{2j}{}^z)^2\rangle_{\text{Av}}][\langle S_{1j}{}^z\rangle_{\text{Av}}+\tfrac{1}{2}]\\ &\times\textstyle\sum_k K_{jk}''(\gamma_{jk}{}^4-\beta_{jk}{}^2\gamma_{jk}{}^2)S_1{}^y\\ &+[S(S+1)-3\langle(S_{1j}{}^z)^2\rangle_{\text{Av}}][\langle S_{2j}{}^z\rangle_{\text{Av}}+\tfrac{1}{2}]\\ &\times\textstyle\sum_k K_{jk}''(-2\beta_{jk}{}^2\gamma_{jk}{}^2)S_2{}^y.\end{aligned}$$

An assumption equivalent to the classical assumption of equal anisotropy fields seen by both sublattices is that in the above equation

$$\langle(S_{1j}{}^z)^2\rangle_{\text{Av}}=\langle(S_{2j}{}^z)^2\rangle_{\text{Av}};\quad [\langle S_{1j}{}^z\rangle_{\text{Av}}+\tfrac{1}{2}]=-[\langle S_{2j}{}^z\rangle_{\text{Av}}+\tfrac{1}{2}];$$

and now defining

$$\gamma K_1/M=\hbar^{-1}q\textstyle\sum_k K_{jk}(\gamma_{jk}{}^4-3\beta_{jk}{}^2\gamma_{jk}{}^2); \tag{50a}$$

$$\gamma K_2/M=\hbar^{-1}q\textstyle\sum_k K_{jk}''(\gamma_{jk}{}^4-\beta_{jk}{}^2\gamma_{jk}{}^2); \tag{50b}$$

$$\gamma K_3/M=\hbar^{-1}q\textstyle\sum_k K_{jk}''(2\beta_{jk}{}^2\gamma_{jk}{}^2); \tag{50c}$$

[13] J. H. Van Vleck, Phys. Rev. 74, 1168 (1948).

[14] Strictly, these will not drop out unless we have cubic symmetry within each of the two magnetic sublattices considered. Shull's experiments (see reference 6) indicate there is no such symmetry in, e.g., MnO. However, there is, of course, symmetry within each of the eight actual sublattices of a f.c.c., and if we make the reasonable assumption that the anisotropy is associated with the super-exchange, the terms in question will go to zero.

where

$$q=-[S(S+1)-3\langle(S_{1j}{}^{x})^{2}\rangle_{\mathrm{Av}}][\langle S_{1j}{}^{x}\rangle_{\mathrm{Av}}+\tfrac{1}{2}],$$

we have

$$Q=-\gamma[(K_1+K_2)/M]S_1{}^{y}-\gamma(K_2/M)S_2{}^{y}.$$

Thus Eq. (46) has become

$$dS_1{}^{x}/dt=-\gamma[H_0-\lambda M_2{}^{z}+(K_1+K_2)/M]S_1{}^{y}$$
$$-\gamma[\lambda M_1{}^{z}+(K_2/M)]S_2{}^{y}; \quad (51)$$

with similar equations in $S_2{}^{x}$, etc. Equation (51) is the same as Eq. (2) (in absence of rf field $\mathbf{H}$) when account is taken of the anisotropy refinements of Sec. IV A. Thus we have shown that the quantum-mechanical and classical approaches lead to the same differential equations of motion, and hence to the same resonance frequencies.

It is noted that for a simple cubic lattice Eq. (50c) gives $K_2/M=0$. For more complicated lattices Eq. (50) ought not to be taken too seriously, as witness reference 14.

VII. EXPERIMENTAL SITUATION

A sudden and almost complete extinction of paramagnetic resonance absorption at microwave frequencies on cooling an antiferromagnetic below the Curie point has been noticed by several investigators. Using a single crystal of the tetragonal salt MnF_2 ($T_c\simeq 70°K$) in a field $H_0\simeq 3000$ oersteds, Hutchison[15] has found the absorption to vanish at 64°K, independent of the orientation of H_0. The same effect has been noticed by Maxwell *et al.*,[16] using sintered Cr_2O_3 ($T_c\simeq 311°K$) in a field $H_0=3450$ oersteds.

On the other hand a continuation of strong paramagnetic absorption well below the Curie point in thin disk polycrystalline samples of the cubic compounds MnS, MnO, and MnSe has been reported from Japan.[17] In particular as MnS ($T_c\simeq 198°K$) was cooled to 78°K the resonance line at 3510 oersteds decreased quite smoothly in amplitude and increased in width. A similar result, but with the intensity more sharply reduced below the Curie point, has been found in powdered Cr_2O_3 by Maxwell *et al.*[16] The latter experiment also showed a second line below the Curie point appearing at a smaller and smaller value of H_0 (fixed ω) as the temperature decreased. It may be that these lines are caused by the presence of ferromagnetic impurities in the lattice.

The antiferromagnetic resonance discussed in this paper probably would not have been observed in any of the published experiments, because they were carried out rather far from the expected fields and frequencies. Antiferromagnetic resonance has apparently been observed at Leiden[2] in the experiment with $CuCl_2\cdot 2H_2O$ mentioned in Sec. I. It would be of interest if experiments could be done with other crystals of very low Curie points, or with some of the above-mentioned crystals in fields $H_0\sim 5\times 10^4$ oersteds.

ACKNOWLEDGMENTS

This research has been assisted in part by the ONR. Part of it was done while one of us (F. K.) held a Whiting Fellowship. We wish to thank Mr. H. Kaplan for checking the derivations. We also are very indebted to Professor C. A. Hutchison and Dr. L. R. Maxwell for several communications regarding their experiments.

[15] C. A. Hutchison (private communication).

[16] Trounson, Bleil, Wangsness, and Maxwell, Phys. Rev. **79**, 542 (1950).

[17] Okamura, Torizuka, and Kojima, Phys. Rev. **82**, 285 (1951).

A Collective Description of Electron Interactions: II. Collective *vs* Individual Particle Aspects of the Interactions

DAVID PINES
Randal Morgan Laboratory of Physics, University of Pennsylvania, Philadelphia, Pennsylvania

AND

DAVID BOHM*
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey

(Received September 28, 1951)

The behavior of the electrons in a dense electron gas is analyzed in terms of their density fluctuations. These density fluctuations may be split into two components. One component is associated with the organized oscillation of the system as a whole, the so-called "plasma" oscillation. The other is associated with the random thermal motion of the individual electrons and shows no collective behavior. It represents a collection of individual electrons surrounded by comoving clouds of charge which screen the electron fields within a distance of the order of magnitude of the Debye length. This split up of the density fluctuations corresponds to an effective separation of the Coulomb interaction into long-range and short-range parts; the separation occurs at roughly the Debye length.

The relation between the individual and collective aspects of the electron gas is discussed in detail, and a general physical picture of the behavior of the system is given. It is shown that for phenomena involving distances greater than the Debye length, the system behaves collectively; for distances shorter than this length, it may be treated as a collection of approximately free individual particles, whose interactions may be described in terms of two-body collisions.

This approach is used to study the interaction of a specified electron with the remainder of the electron gas. It is shown that the collective part of the response of this remainder to the field of the specified particle screens this field within a distance of the order of the Debye length; this furnishes a detailed description of the screening process. Moreover, if the specified particle moves with greater than the mean thermal speed, it excites collective oscillations in the form of a wake trailing the particle. The frequency of these collective oscillations and the energy emitted by the particle are calculated. A correspondence theoretical method is used to treat this phenomenon for the electrons in a metal. The results are in good agreement with the experiments of Ruthemann and Lang on the energy loss of kilovolt electrons in this metallic films.

The generalization of these methods to an arbitrary interparticle force is carried out, and a criterion is obtained for the validity of a collective description of the particle interactions. It is shown that strong forces and high particle density favor collective behavior, while high random thermal velocities oppose it.

I. INTRODUCTION

IN this paper we wish to develop a detailed physical picture of the behavior of the electrons in a dense electron gas. We do this with the aid of a collective description of the particle motion. In a previous paper,[1] hereafter referred to as I, we used a collective description in treating the organized behavior of the electrons resulting from the transverse electromagnetic interactions. This was done by means of a canonical transformation to a set of collective coordinates which were appropriate for a description of the organized behavior. In the present paper, we are concerned with the organization produced by the Coulomb interactions, which are far more important quantitatively than the transverse electromagnetic interactions. We stress the physical picture of the electron behavior here, because it is essential for the proper development and understanding of the necessary mathematical formulation. In a subsequent paper we shall extend our results to the quantum theory by developing the canonical transformation for the Coulomb case in a manner similar to that given in I.

In a dense electron gas, the particles interact strongly because of the long range of the Coulomb force; in fact, each particle interacts simultaneously with all the other particles. As a result the equations of motion become extremely difficult to solve. The usual perturbation theory solution based on the assumption of a small interaction between pairs of particles breaks down. As was the case in I, a collective description provides a far better starting point for a solution than a description in terms of the individual particles. For the collective description makes possible a simple method of treating the simultaneous interaction of many electrons (as opposed to the individual particles approach which gives a simple method of treating two-body collisions).

Certain examples of collective behavior in an electron gas are well known from the study of gaseous discharges. These are the organized oscillations of the system as a whole, the "plasma" oscillations.[2,3] These oscillations have been studied theoretically with the simplifying assumption that the gas is composed of a distribution of beams of charge, each beam having a well-defined velocity at each point in space.[4,5] This approach, although it gives many useful and instructive results concerning the oscillations, represents an excessive abstraction which is not capable of describing many other important aspects of the organized behavior in the gas.

* Now at Physics Department, University of Sao Paulo, Sao Paulo, Brazil.

[1] D. Bohm and D. Pines, Phys. Rev. **82**, 625 (1951).

[2] L. Tonks and I. Langmuir, Phys. Rev. **33**, 195 (1929).

[3] H. J. Merrill and H. W. Webb, Phys. Rev. **55**, 1191 (1939).

[4] A. Vlasov, J. Phys. (U.S.S.R.) **9**, 25, 130 (1945).

[5] D. Bohm and E. P. Gross, Phys. Rev. **75**, 1851 and 1864 (1949); Paper A discusses the origin of medium-like behavior, and gives many references to the earlier work on plasma oscillations; Paper B deals with the excitation and damping of oscillations.

In this paper we start with a gas of interacting point electrons. Instead of following the motion of the individual particles, we describe the gas in terms of the Fourier components of the electron density at each point in space. These Fourier components are proportional to the density fluctuations in the electron gas. We find that the density fluctuations can be split into two parts. One part represents an organized oscillation with the characteristic "plasma" frequency, and is clearly associated with the collective behavior of the system. The other part is associated with the random thermal motion of the individual particles, and shows no collective behavior. For wavelengths greater than a certain critical length λ_D (the Debye length), the fluctuations are primarily collective. When this is the case a description of the behavior of the electrons solely in terms of their collective motion is a good approximation. For wavelengths smaller than λ_D, however, the fluctuations are primarily associated with individual particle motion, and in this case the system is best described by following the individual particles.

We use the above split-up of the density fluctuations to study the collective response of the electron gas to the field of an individual charged particle moving with a specified velocity, $\mathbf{v}_0$. When $\mathbf{v}_0$ is less than the mean thermal speed of the gas, we find that the collective response is just such as to screen out the field of the specified particle within a distance of order of λ_D. When $\mathbf{v}_0$ is greater than the mean thermal speed, a similar screening occurs, but in addition a new phenomenon appears: *viz.*, the excitation of collective oscillations in the form of a wake trailing the particle. The phenomenon resembles the Čerenkov radiation produced by fast electrons in dielectric materials. Experiments by Ruthemann and Lang, on the bombardment of thin metallic films by fast electrons tend to verify our theoretical predictions concerning this type of excitation of collective oscillations.[6,7]

On the basis of the above results we are led to the following physical picture of the screening process. As any electron moves through the assembly, the other electrons are pushed away from it by the Coulomb repulsion. Each particle is thus surrounded by a cloud of extent λ_D, in which there is a deficiency of electrons, which is responsible for screening the field of the particle in question. As a result of this screening, the cross section for interparticle collision is so greatly reduced that the mean-free path of an electron is considerably greater than the interparticle spacing. Thus for many purposes the electron plus its associated cloud may be regarded as an effective free particle.

The splitup of the density fluctuation into collective and individual particles components may be viewed in the following way. The collective part includes the effects of the long range of the Coulomb force which leads to the simultaneous interaction of many particles. The individual particles component represents the density fluctuations arising from the randomly moving individual particles plus their comoving electron clouds, and thus includes the effects of the residual short-range screened Coulomb force, which leads only to two-body collisions.

Our main conclusion is that neither the collective description nor the individual particles description of the electron gas is by itself entirely adequate. For not only in each description needed in its appropriate region, but also the interaction between collective and individual aspects determines many important properties of the system. It is just this synthesis of individual and collective aspects that makes the electron gas such an exceptionally interesting medium.

There are a wide variety of systems to which the methods developed in this paper may be applied. Some of these are: (a) Ion gases of high density (plasmas), which occur in gaseous discharges, interstellar nebulae, atmosphere of sun and stars, ionosphere of earth, etc. (b) The electrons and ions in a metal. The ions in a metal are also susceptible to a collective description, and, in interaction with electrons, they give rise to sound waves, whose properties can be calculated with the collective method. In this way, one can obtain an improved treatment of the so-called "lattice-electron" interaction, which is important in the theory of electrical conductivity, and probably in superconductivity.[8]

In Appendix II, we discuss the generalization of these methods to an arbitrary interparticle force. Preliminary considerations indicate that the collective description may be applicable to the particles in an atomic nucleus.[9]

II. COLLECTIVE OSCILLATION

We begin by a study of the way in which the interactions in an assembly of electrons bring about organized behavior and collective oscillation. We shall consider an aggregate of approximately free electrons embedded in a medium of fixed positive charges whose average density is equal to that of the electrons. For most purposes, this distribution of charge can be regarded as uniformly smeared out throughout the entire system. Hence, it merely serves to neutralize the mean electron charge. The previous simplifications are adequate for the treatment of electron behavior in gaseous discharges, as well as in those metals in which the effects of the lattice periodicity are not very important. The approximations used here will also provide a good starting point for the investigation of other metals.

Each electron in the assembly is acted on by the sum of the forces arising from all of the other electrons plus

[6] G. Ruthemann, Ann phys. **2**, 113 (1948).
[7] W. Lang, Optik **3**, 233 (1948).

[8] This problem is now under investigation by one of us (D.B.) and T. Staver. Preliminary results are given in Phys. Rev. **84**, 836 (1951).
[9] One of us (D.P.) and M. I. Ferentz are currently investigating a nuclear model based on a collective description of nucleon interactions.

that resulting from the smeared-out positive charge. The potential energy of interaction between the ith and jth electrons, $e^2/|\mathbf{x}_i-\mathbf{x}_j|$, may be expanded as a Fourier series in a box of unit volume with periodic boundary conditions, and is

$$(e^2/|\mathbf{x}_i-\mathbf{x}_j|)=4\pi e^2\sum_k(1/k^2)e^{i\mathbf{k}\cdot(\mathbf{x}_i-\mathbf{x}_j)}. \quad (1)$$

The equation of motion of the ith electron is given by

$$\ddot{\mathbf{x}}_i=-(4\pi e^2 i/m)\sum'_{j,k}(\mathbf{k}/k^2)e^{i\mathbf{k}\cdot(\mathbf{x}_i-\mathbf{x}_j)}, \quad (2)$$

where the prime denotes a sum in which $\mathbf{k}=0$ is excluded.[10] (In the previous sum, the term with $j=i$ should also be excluded. However, since $1/k^2$ is a spherically symmetric function of $\mathbf{k}$, this term averages out to zero and therefore does not have to be explicitly excluded.)

Equation (2) is in general extremely difficult to solve by following the motions of the individual particles, especially when the assembly is so dense that many-body collisions become important. Moreover, the range of the Coulomb potential is so great that many-body collisions are important even in an electron gas of low density. Under these conditions, the electrons move together in organized fashion, and one finds the well-known phenomenon of "plasma" oscillations of the system as a whole.[2] In these cases the usual method of approximation in solving (2) which involves the assumption that the interaction forces produce small perturbations on uniform straight line particle motion, fails. Our approach to the equations of motion is aimed at making use of the simplicity of the collective behavior as a starting point for a tractable solution. We shall see that this method leads to a far better description of the motion than one based solely on following the individual particles.

We shall be primarily concerned with two questions: the description of the motion of the particle assembly in terms of its collective behavior, and the limits on the applicability of such a collective description. As a first step, we study the fluctuations in the particle density, because, as we shall see, their behavior provides a good measure of the applicability of a collective description. We assume, for all practical purposes, that we are dealing with point particles, so that the particle density in our box of unit volume is given by,

$$\rho(\mathbf{x})=\sum_i\delta(\mathbf{x}-\mathbf{x}_i). \quad (3)$$

We shall find it more convenient to work with the Fourier components, ρ_k, of the density. These are given by,

$$\rho_k=\int d\mathbf{x}\rho(\mathbf{x})e^{-i\mathbf{k}\cdot\mathbf{x}}=\sum_i e^{-i\mathbf{k}\cdot\mathbf{x}_i}, \quad (4)$$

and

$$\rho(\mathbf{x})=\sum_k\rho_k e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_i)}. \quad (5)$$

We note that ρ_0 represents the mean electron density, n, and the ρ_k with $k\neq 0$ describe fluctuations about that mean density. It is readily verified that the equations of motion (2), may be re-expressed as

$$\ddot{x}_i=-(4\pi e^2 i/m)\sum'_k(\mathbf{k}/k^2)\rho_k e^{i\mathbf{k}\cdot\mathbf{x}_i}. \quad (6)$$

The ρ_k thus determine the force acting on each particle.

We now obtain the equations describing the time behavior of the ρ_k. On differentiating (4), we have

$$\dot{\rho}_k=-i\sum_i(\mathbf{k}\cdot\mathbf{v}_i)e^{-i\mathbf{k}\cdot\mathbf{x}_i}, \quad (7a)$$

$$d^2\rho_k/dt^2=-\sum_i[(\mathbf{k}\cdot\mathbf{v}_i)^2+i\mathbf{k}\cdot\dot{\mathbf{v}}_i]e^{-i\mathbf{k}\cdot\mathbf{x}_i}. \quad (7b)$$

We obtain $\dot{\mathbf{v}}_i$ from the equations of motion (2), and $d^2\rho_k/dt^2$ becomes

$$d^2\rho_k/dt^2=-\sum_i(\mathbf{k}\cdot\mathbf{v}_i)^2e^{-i\mathbf{k}\cdot\mathbf{x}_i}-\sum_{\substack{k'ij\\ k'\neq 0}}[4\pi e^2/m(k')^2]\mathbf{k}\cdot\mathbf{k}'\{\exp[i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{x}_i]\}\exp(-i\mathbf{k}'\cdot\mathbf{x}_j). \quad (8)$$

It is in the treatment of the second term on the right-hand side of the previous equation that we encounter one of the central approximations used in the development of a collective description. For we split the sum over $\mathbf{k}'$ into two parts. The first part, with $\mathbf{k}'=\mathbf{k}$, is independent of the coordinate, $\mathbf{x}_i$, so that the sum over i yields n, the total number of particles.[11] The second part, (those terms with $\mathbf{k}'\neq\mathbf{k}$) contains phase factors $\exp i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{x}_i$, which depend on the positions of the particles. These terms tend to average out to zero, since there are a very large number of particles distributed very nearly in random positions. As a first approximation, we neglect such terms. This procedure we call the random phase approximation. The validity of this approximation will be demonstrated in detail in Sec. VI. Using the random phase approximation, we then obtain

$$d^2\rho_k/dt^2=-\sum_i(\mathbf{k}\cdot\mathbf{v}_i)^2e^{-i\mathbf{k}\cdot\mathbf{x}_i}-(4\pi ne^2/m)\sum_i e^{-i\mathbf{k}\cdot\mathbf{x}_i}. \quad (9)$$

The first term on the right-hand side of (9) is one that would be present even in the absence of particle interaction, and arises simply from the random thermal motion of the individual particles. The second term represents the effects of particle interactions. For sufficiently small $\mathbf{k}$ it is clear that the first term can be neglected in comparison with the second. Under these conditions, (9) becomes

$$d^2\rho_k/dt^2+(4\pi ne^2/m)\rho_k=0. \quad (10)$$

[10] The exclusion of the term with $k=0$ takes into account the uniform background of positive charge, and hence the over-all charge neutrality of the system. This may readily be seen by Fourier-analyzing the electron charge density and noting that Fourier-component corresponding to $k=0$ describes the mean density of electrons which is canceled by the equal density of uniformly distributed positive charge. The Fourier-expansion for the potential at a given point there fore does not contain the term with $k=0$.

[11] Since we are working with a unit volume, n is numerically equal to the mean density.

Thus, as a result of the Coulomb interaction, the electron density oscillates with the well-known "plasma" frequency,

$$\omega_P=(4\pi ne^2/m)^{\frac{1}{2}}. \tag{11}$$

The excitation of a particular ρ_k corresponds to a wave-like density fluctuation, analogous to a sound wave.

Let us now consider the physical significance of the terms which determine the time variation of the ρ_k. In the absence of interaction, each particle moves in a straight line with a constant velocity, $\mathbf{v}_{0i}$. In this case $\mathbf{x}_i=\mathbf{x}_{0i}+\mathbf{v}_{0i}t$, and ρ_k takes the form

$$\rho_k{}^0=\textstyle\sum_i e^{i\mathbf{k}\cdot(\mathbf{x}_{0i}+\mathbf{v}_{0i}t)}. \tag{12}$$

In order that $\rho_k{}^0$ be large at a particular time, say $t=0$, it is necessary that the $\mathbf{x}_{0i}$ be so distributed that the various terms $e^{i\mathbf{k}\cdot\mathbf{x}_{0i}}$ arising from each particle tend to be in phase. However, each of these terms oscillates with an angular frequency $\mathbf{k}\cdot\mathbf{v}_{0i}$, which is, in general, different from that of any other. Consequently, even if they initially have definite phase relations, the contribution of different particles to ρ_k will soon tend to get out of phase and cancel each other. This means that in a gas of free particles a particular fluctuation cannot persist for an appreciable length of time. The tendency of the random motion of each particle to contribute a term to the density variation having an angular frequency, $\omega=\mathbf{k}\cdot\mathbf{v}_i$, is reflected in the first term on the right-hand side of (9), which by itself would result in a similar type of time variation of the ρ_k. We conclude then that a collection of free particles shows no organized behavior, but that instead its characteristic property is that a disturbance tends to die out as a result of the random diffusion of the particles. On the other hand, the effect of the Coulomb force of interaction, appearing in the second term on the right-hand side of (9), is to cause each particle to make a contribution to $d^2\rho_k/dt^2$ which oscillates with the same angular frequency as that of every other particle. Thus, if the random thermal motions were not present the Coulomb forces would produce perfectly organized behavior of the ρ_k. Actually, of course, both the Coulomb forces and the random thermal motions are present simultaneously, so that the net behavior of the electron gas will show some collective aspects and some of the aspects of an aggregate of randomly moving individual particles. To the extent that the effects of the Coulomb forces predominate in determining $d^2\rho/dt^2$, the gas will display primarily its collective aspect because the ρ_k determine the force on each particle and therefore the behavior of the entire system. On the other hand, when the random individual particle motions predominate in determining $d^2\rho_k/dt^2$, the collective description will be inappropriate, and a description of the system in terms of the motions of the individual particles becomes a better starting point. A rough criterion for the applicability of a collective description is therefore that for most particles the collective term in (9) be much greater than the term arising from the random thermal individual particle motions, or that

$$(4\pi ne^2/m)\gg\langle(\mathbf{k}\cdot\mathbf{v}_i)^2\rangle_{\text{Av}}. \tag{13}$$

Thus we see that a high particle density favors organized oscillation, while high random velocities oppose it. A strong force of interaction (as measured by the e^2 factor) also favors organized behavior.

For an electron gas with an isotropic Maxwellian velocity distribution at a temperature T, our criterion becomes

$$k^2\ll(12\pi ne^2/m\langle v_i{}^2\rangle_{\text{Av}})=(4\pi ne^2/\kappa T)=\lambda_D{}^{-2}, \tag{14}$$

where κ is Boltzmann's constant, and λ_D is the well-known Debye length,[12] defined by

$$\lambda_D=(\kappa T/4\pi ne^2)=(1/3)[\langle\langle v_i{}^2\rangle_{\text{Av}}\rangle/\omega_P{}^2]. \tag{15}$$

This means that in a collection of electrons, organized behavior is most important in phenomena connected with distances greater than λ_D, while for phenomena connected with shorter distances the individual particles point of view provides the best starting point. This result has already been obtained by other investigators using different methods.[13] It is because of the long range of the Coulomb force (as expressed in terms of the $1/k^2$ in its Fourier component) that the collective approximation is sure to apply if one goes to large enough wavelengths ($\lambda>\lambda_D$). A long-range force favors the collective approximation at long wavelengths, because a large number of particles can be made to contribute cumulatively to a collective oscillation in the charge density.

It is of interest to evaluate λ_D for a few examples. Thus in a typical gaseous discharge of density $n=10^{12}$, and a mean electron kinetic energy of 3 ev, we obtain $\omega_P=5\times10^{10}$ and $\lambda_D\cong10^{-3}$ cm. For an electron gas density $\sim10^{23}$, such as one finds in a typical metal, $\omega_P\cong2\times10^{16}$, and $\lambda_D\cong0.4$A, as compared with an interparticle spacing of ~2A.[14]

We have seen that the time variation of the ρ_k reflects both the collective and the individual particle aspects of the assembly, so that the ρ_k do not oscillate harmonically for appreciable values of k, and therefore do not consistitue a satisfactory set of collective coordinates. We might expect that somewhat different functions of the particle coordinates and velocities exist which do oscillate harmonically for nonzero k, despite the effects of random thermal motion. For small

[12] The Debye length was first introduced in connection with screening processes in highly ionized electrolytes. P. Debye and E. Huckel, Physik Z. 24, 185 (1923).

[13] See D. Bohm and E. P. Gross, Paper A, reference 5.

[14] For this case λ_D clearly cannot constitute the minimum distance at which organization becomes important, since it is less than the interparticle spacing. However, this indicates that the organization is very important for the electrons in a metal, since it does extend down to regions of the order of the interparticle spacing. The appropriate minimum distance for a degenerate electron gas will be discussed in our subsequent paper on the quantum theory.

k, these functions should approach ρ_k. We now demonstrate that such functions exist and are in fact proportional to

$$q_k=\sum_i\{1/[\omega^2-(\mathbf{k}\cdot\mathbf{v}_i)^2]\}e^{-i\mathbf{k}\cdot\mathbf{x}_i}. \quad (16)$$

We shall find it convenient to introduce the quantities

$$\xi_{k,\omega}=\sum_i[1/(\omega-\mathbf{k}\cdot\mathbf{v}_i)]e^{-i\mathbf{k}\cdot\mathbf{x}_i}, \quad (17)$$

in terms of which the q_k may be expressed as

$$q_k=(1/2)[\xi_{k,\omega}-\xi_{k,-\omega})/\omega]. \quad (18)$$

We shall show that the q_k satisfy the harmonic oscillator equation

$$\ddot{q}_k+\omega^2 q_k=0, \quad (19)$$

by proving that

$$\dot{\xi}_{k,\omega}+i\omega\xi_{k,\omega}=0. \quad (20)$$

That (19) is a consequence of (20) may be readily verified by differentiating (18).

To prove that the $\xi_{k,\omega}$ satisfy (20), we differentiate (17), obtaining

$$\dot{\xi}_{k,\omega}=-i\sum_i\{[(\mathbf{k}\cdot\mathbf{v}_i)/(\omega-\mathbf{k}\cdot\mathbf{v}_i)] +(\mathbf{k}\cdot\dot{\mathbf{v}}_i)/(\omega-\mathbf{k}\cdot\mathbf{v}_i)^2\}e^{-i\mathbf{k}\cdot\mathbf{x}_i}. \quad (21)$$

Using the equations of motion (2), we find:

$$\dot{\xi}_{k,\omega}+i\omega\xi_{k,\omega}=i\sum_i e^{-i\mathbf{k}\cdot\mathbf{x}_i}-i\sum_{\mathbf{k}'}\frac{(\mathbf{k}\cdot\mathbf{k}')4\pi e^2}{mk^2(\omega-\mathbf{k}\cdot\mathbf{v}_i)^2} \times[\exp i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{x}_i]\exp(-i\mathbf{k}'\cdot\mathbf{x}_j). \quad (22)$$

As in Eq. (8), the terms with $\mathbf{k}'\neq\mathbf{k}$ in the sum on the right hand side will be neglected in the random phase approximation. Retaining only the terms with $\mathbf{k}'=\mathbf{k}$, we then obtain for the right-hand side, after interchanging i and j.

$$i\sum e^{-i\mathbf{k}\cdot\mathbf{x}_i}-i\sum_{ij}\frac{4\pi e^2}{m(\omega-\mathbf{k}\cdot\mathbf{v}_j)^2}e^{-i\mathbf{k}\cdot\mathbf{x}_i}.$$

In order that $\xi_{k,\omega}$ shall oscillate harmonically, the previous expression must vanish for abitrary $\mathbf{x}_i$. Thus we find that the $\xi_{k,\omega}$, and hence the q_k, oscillate harmonically, provided ω satisfies the following dispersion relation

$$1=(4\pi e^2/m)\sum_j 1/(\omega-\mathbf{k}\cdot\mathbf{v}_j)^2. \quad (23)$$

This same dispersion relation has already been obtained by other investigators using the beams treatment.[15] It is an integral equation which is difficult to solve exactly. However, for sufficiently small k, the denominator in (23) may be expanded in a series of powers of $(\mathbf{k}\cdot\mathbf{v}_j)/\omega$, and we obtain the approximate dispersion relation

$$\omega^2=\omega_P^2+k^2\langle v^2\rangle_{Av}, \quad (24)$$

provided we assume an isotropic distribution of velocities.[16] We note that for small k this reduced to our earlier dispersion relation, (11). The condition for the validity of (24) is essentially our criterion for the applicability of the collective description, (14). We also note that for $[(\mathbf{k}\cdot\mathbf{v})/\omega]\ll 1$, q_k approaches proportionality to ρ_k.

Where the described expansion is not very accurate, ω can be obtained by a numerical solution of (23). Other investigators have shown that no solution exists for k larger than a critical value of the order of λ_D^{-1}.[15] This result confirms our general physical picture in which the collective description is to be used only for distances $>\lambda_D$. Furthermore, over almost all values of k for which (23) has a solution, the approximate dispersion relation (24) is valid.

III. SEPARATION BETWEEN COLLECTIVE AND INDIVIDUAL COMPONENTS OF DENSITY FLUCTUATIONS

We have seen that the density fluctuations, ρ_k, can, for sufficiently small k, be described almost entirely in terms of the collective coordinate, q_k. On the other hand, since there are no collective coordinates for $k\gtrsim\lambda_D^{-1}$, the density fluctuations here must be associated primarily with the random thermal motions of the individual particles. For intermediate values of k, both types of fluctuations may be expected to be significant. In this section, we shall treat the problem of separating individual and collective aspects of a general fluctuation by splitting ρ_k into two parts as follows,

$$\rho_k=a_k q_k+\eta_k, \quad (25)$$

where a_k is a suitable constant, which will be chosen by methods to be developed presently. The first part, $a_k q_k$, is clearly a collectively describable part of ρ_k. We shall see that a choice of a_k is possible such that η_k describes fluctuations associated only with the random thermal motions of the individual particles. Moreover, the η_k will be shown to be the Fourier coefficients of a density distribution in which each electron is surrounded by a comoving cloud containing a deficiency of electrons, which screens the field of the given electron within a distance of the order of λ_D.

We begin with a consideration of (25) for $k\lesssim k_D$, where $k_D\cong(1/\lambda_D)$. Using our definition (16) of q_k, we may express η_k as

$$\eta_k=\sum_i\{1-a_k/[\omega^2-(\mathbf{k}\cdot\mathbf{v}_i)^2]\}e^{-i\mathbf{k}\cdot\mathbf{x}_i}. \quad (26)$$

[15] For a more detailed discussion of the dispersion relation (Eq. (23)] see Bohm and Gross, Paper A, reference 5.

[16] The mean kinetic energy will depend both on the temperature, which determines the random thermal motion, and on the amplitude of organized oscillation. Normally, the thermal energy is much greater than the organized oscillation energy, so that for all practical purposes, the frequency does not depend appreciably on the amplitude of organized oscillation. The slight dependence on amplitude resembles a similar phenomenon obtained in connection with sound waves, which likewise increase in frequency as their amplitude increases. In any case, the entire effect is in the domain of the nonlinear aspects of the problem, and therefore can be neglected in a linear approximation. Within the linear approximation, the $\langle v^2\rangle_{Av}$ appearing in the dispersion relation should be the value existing in the absence of organized oscillation.

With an arbitrary choice of a_k, $\ddot{\eta}_k$, like $d^2\rho_k/dt^2$ in Eq. (9), may be expected to have two kinds of terms. The first kind are those we obtain in the absence of interparticle forces, (in which case $\dot{v}_i=0$), and are

$$\ddot{\eta}_k^{(1)}=\sum_i-(\mathbf{k}\cdot\mathbf{v}_i)^2\{1-[a_k/[\omega^2-(\mathbf{k}\cdot\mathbf{v}_i)^2]]\}e^{-i\mathbf{k}\cdot\mathbf{x}_i}. \quad (27)$$

These arise from the random thermal motion of the particles, and can in no way possess an organized component. The second kind arise because of the interparticle forces, and, like the second term on the right-hand side of (9), represent the organized behavior of the system. To insure that η_k has no collective component, we must choose a_k in such a way that terms of this second kind vanish and $\ddot{\eta}_k$ is given by (27).

Using Eqs. (9) and (19), we obtain

$$\begin{aligned}d^2\eta_k/dt^2&=d^2\rho_k/dt^2-a_kd^2q_k/dt^2\\&=\sum_i\{-(\mathbf{k}\cdot\mathbf{v}_i)^2-\omega_P^2\\&\quad+a_k\{\omega^2/[\omega^2-(\mathbf{k}\cdot\mathbf{v}_i)^2]\}\}e^{-i\mathbf{k}\cdot\mathbf{x}_i},\end{aligned} \quad (28)$$

where we have continued to use the random phase approximation. This may be rearranged as

$$\ddot{\eta}_k=\ddot{\eta}_k^{(1)}+\sum_i\{(a_k-\omega_P^2)/[\omega^2-(\mathbf{k}\cdot\mathbf{v}_i)^2]\}e^{-i\mathbf{k}\cdot\mathbf{x}_i}. \quad (29)$$

Hence, if we choose $a_k=\omega_P^2$, we obtain the desired separation with

$$\eta_k=\sum_i\frac{[\omega^2-\omega_P^2-(\mathbf{k}\cdot\mathbf{v}_i)^2]}{\omega^2-(\mathbf{k}\cdot\mathbf{v}_i)^2}e^{-i\mathbf{k}\cdot\mathbf{x}_i}. \quad (30)$$

It will be convenient to include the coefficient ω_P^2 in our basic definition of q_k so that we have

$$q_k=\sum_i\{\omega_P^2/[\omega^2-(\mathbf{k}\cdot\mathbf{v}_i)^2]\}e^{-i\mathbf{k}\cdot\mathbf{x}_i}, \quad (31)$$

and

$$\rho_k=q_k+\eta_k \quad (k<k_D). \quad (32)$$

With these definitions, we see that the charge density has split into two independent parts. One of these, q_k, oscillates harmonically with frequency ω, and can be given an arbitrary amplitude of oscillation dependent on the boundary conditions. The other part, η_k, has no collective behavior, and is therefore best regarded simply as the sum of independent contributions from individual particles. This part is present even in the absence of organized oscillation.

When $k\gtrsim k_D$, there are no collective coordinates, so that $\rho_k=\eta_k$. However, ρ_k, (or η_k), in this region does not satisfy an equation of motion like (27) since we still have the second term on the right-hand side of (9), *viz.*, $\omega_P^2\rho_k$, in the expression for $\ddot{\eta}_k$. However, for values of k appreciably greater than k_D this term is negligible, and hence, except for a region of $k\sim k_D$, which we shall see is of no great importance in most problems, we may regard η_k as satisfying (27) for all values of k.

We now wish to investigate in detail the meaning of the density fluctuations. To do this, we consider the contribution to the spatial distribution of electron density arising from that part of η_k associated with the rth particle,

$$\begin{aligned}\eta_r(\mathbf{x})&=\sum_k\eta_{kr}e^{-i(\mathbf{k}\cdot\mathbf{x})}\\&=\sum_{k<k_D}\left\{\frac{\omega^2-\omega_P^2-(\mathbf{k}\cdot\mathbf{v}_r)^2}{\omega^2-(\mathbf{k}\cdot\mathbf{v}_r)^2}\right\}e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_r)}\\&\quad+\sum_{k>k_D}e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_r)}.\end{aligned} \quad (33)$$

In the first term on the right-hand side, since we are limited to $k<k_D$, we may expand ω^2 according to (24), obtaining,

$$\begin{aligned}\eta_r(\mathbf{x})&=\sum_{k<k_D}\left\{\frac{k^2\langle v^2\rangle_{kr}-(\mathbf{k}\cdot\mathbf{v}_r)^2}{\omega_P^2+k^2\langle v^2\rangle_{kr}-(\mathbf{k}\cdot\mathbf{v}_r)^2}\right\}e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_r)}\\&\quad+\sum_{k>k_D}e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_r)}.\end{aligned} \quad (34)$$

We note that for $k\gg k_D$,

$$k^2\langle v^2\rangle_{kr}-(\mathbf{k}\cdot\mathbf{v}_r)^2\cong\omega_P^2+k^2\langle v^2\rangle_{kr}-(\mathbf{k}\cdot\mathbf{v}_r)^2.$$

Hence, as an approximation, we can take

$$\eta_r(\mathbf{x})\cong\sum_{\text{all }k}\left\{\frac{k^2\langle v^2\rangle_{kr}-(\mathbf{k}\cdot\mathbf{v}_r)^2}{\omega_P^2+k^2\langle v^2\rangle_{kr}-(\mathbf{k}\cdot\mathbf{v}_r)^2}\right\}e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_r)}. \quad (35)$$

This is an accurate description of $\eta_r(x)$ for $k\gg k_D$ and $k\ll k_D$. It is not far off in the region $k\cong k_D$.

To evaluate the sum over $\mathbf{k}$ we choose the direction of $\mathbf{v}_r$ as the z axis in $\mathbf{k}$ space. We also make the following change of variables in $\mathbf{k}$ and $\mathbf{x}$ space

$$\begin{aligned}(\langle v^2\rangle_{kr}-v_r^2)k_z^2&=\langle v^2\rangle_{kr}(k_z')^2,\\\langle v^2\rangle_{kr}(z-z_r)^2&=(\langle v^2\rangle_{kr}-v_r^2)(z'-z_r')^2,\end{aligned}$$

together with $k_x=k_x'$, $k_y=k_y'$, $x=x'$, and $y=y'$. We then obtain, on changing the sum over $\mathbf{k}$ to an integral,

$$\begin{aligned}\eta_r(\mathbf{x})=\frac{1}{(2\pi)^3}&\left\{\frac{\langle v^2\rangle_{kr}}{\langle v^2\rangle_{kr}-v_r^2}\right\}^{\frac{1}{2}}\int d\mathbf{k}'\left\{\frac{(k')^2\langle v^2\rangle_{kr}}{\omega_P^2+(k')^2\langle v^2\rangle_{kr}}\right\}\\&\times\exp[i\mathbf{k}'\cdot(\mathbf{x}'-\mathbf{x}_r')].\end{aligned} \quad (36)$$

This may be re-expressed as

$$\begin{aligned}\eta_r(\mathbf{x})=-\left(\frac{1}{2\pi}\right)^3&\left\{\frac{\langle v^2\rangle_{kr}}{\langle v^2\rangle_{kr}-v_r^2}\right\}^{\frac{1}{2}}\nabla_{x'}{}^2\\&\times\int d\mathbf{k}'\left\{\frac{\exp[i\mathbf{k}'\cdot(\mathbf{x}'-\mathbf{x}_r')]}{(k')^2+(\omega_P{}^2/\langle v^2\rangle_{kr})}\right\}.\end{aligned} \quad (37)$$

The integral over $\mathbf{k}'$ is readily evaluated, and we obtain

$$\begin{aligned}\eta_r(\mathbf{x})=-&\left\{\frac{\langle v^2\rangle_{kr}}{\langle v^2\rangle_{kr}-v_r^2}\right\}^{\frac{1}{2}}\nabla_{x'}{}^2\\&\times\left[\frac{\exp\{-(\omega_P{}^2/\langle v^2\rangle_{kr})^{\frac{1}{2}}|\mathbf{x}'-\mathbf{x}_r'|\}}{4\pi|\mathbf{x}'-\mathbf{x}_r'|}\right].\end{aligned} \quad (38)$$

Using the well-known result, $(\nabla^2-k^2)\{e^{-kr}/r\}=-4\pi\delta(\mathbf{r})$ we get

$$\eta_r(\mathbf{x})=\frac{\langle v^2\rangle_{\text{Av}}^{\frac{1}{2}}}{(\langle v^2\rangle_{\text{Av}}-v_r^2)^{\frac{1}{2}}}\left\{\delta(\mathbf{x}'-\mathbf{x}_r')-\frac{\omega_P^2\exp\{-(\omega_P^2/\langle v^2\rangle_{\text{Av}})^{\frac{1}{2}}|\mathbf{x}'-\mathbf{x}_r'|\}}{4\pi\langle v^2\rangle_{\text{Av}}|\mathbf{x}'-\mathbf{x}_r'|}\right\}. \tag{39}$$

Returning to our original coordinate system, we have:

$$\eta_r(\mathbf{x})=\delta(\mathbf{x}-\mathbf{x}_r)-\frac{\omega_P^2\alpha_r}{4\pi\langle v^2\rangle_{\text{Av}}}\left\{\frac{\exp\{-(\omega_P^2/\langle v^2\rangle_{\text{Av}})^{\frac{1}{2}}[(x-x_r)^2+(y-y_r)^2+\alpha_r^2(z-z_r)^2]^{\frac{1}{2}}\}}{[(x-x_r)^2+(y-y_r)^2+\alpha_r^2(z-z_r)^2]^{\frac{1}{2}}}\right\}, \tag{40}$$

where $\alpha_r^2=\langle v^2\rangle_{\text{Av}}/[\langle v^2\rangle_{\text{Av}}-v_r^2]$.

This represents a particle at $\mathbf{x}=\mathbf{x}_r$ surrounded by a comoving cloud in which the electron density is reduced below the average. The cloud is elliptical in shape, being shortened in the direction of particle motion by the ratio $(\alpha_r)^{-1}$. It should be noted that for those particles with $v_r^2>\langle v^2\rangle_{\text{Av}}$ the present treatment fails. This case will be discussed in Sec. IV. Here we restrict ourselves to those particles moving with less than the mean thermal speed, $[\langle v^2\rangle_{\text{Av}}]^{\frac{1}{2}}$.

The computation of the potential arising from the charge density $-e\eta_r$ is complicated by the elliptical shape of the electron cloud. To get a rough estimate of this potential, we consider the special case of a particle at rest. For this case the cloud is spherical and the potential is found to be

$$-e\frac{\exp\{-|\mathbf{x}-\mathbf{x}_r|(\omega_P^2/\langle v^2\rangle_{\text{Av}})^{\frac{1}{2}}\}}{|\mathbf{x}-\mathbf{x}_r|}.$$

This is a screened Coulomb potential with a screening radius of the order of λ_D.

More generally we can see that screening still takes place in a distance of the order of λ_D when $\mathbf{v}_r$ takes a nonzero value. We recall that the electrons are embedded in a uniform distribution of positive charge. If the negative charge distribution were perfectly uniform, the electric field would vanish. However, the field does not vanish since we have a collection of negative point charges which produce local fluctuations whose Fourier components are the ρ_k. Let us fix our attention on the individual particles component $\eta_r(\mathbf{x})$ of the density associated with the rth particle as given by (40). Without the comoving cloud described by the second term on the right-hand side of (40), the density is a δ-function, which gives rise to a simple Coulomb potential $-e/|\mathbf{x}-\mathbf{x}_r|$. The comoving cloud represents a region from which electrons have been displaced by the repulsive Coulomb potential of the rth electron. As a result this region contains a net positive charge. In fact the total net positive charge may be obtained by integrating our expression for the electron deficiency [as obtained from (40)] over all space, and the result is readily seen to be unity. Since most of this charge is clearly within a region $\sim\lambda_D$ surrounding the electron, we may deduce from Gauss' theorem that the electric field associated with this electron becomes negligible at distances greater than λ_D from the electron. This constitutes a striking demonstration that the individual particles do not produce important effects at distances greater than the Debye length, so that in this region, the collective coordinates q_k provide an adequate description of all important properties of the electron gas. Although this result has thus far obtained only for these particles moving with less than the mean thermal speed, it will be shown in Sec. IV that similar conclusions apply to the remaining particles.

IV. EXCITATION OF COLLECTIVE OSCILLATIONS AND SCREENING

In this section, we study the collective response of the electron gas to an individual particle moving through the system with velocity $\mathbf{v}_0$. We first show that if $\mathbf{v}_0$ is less than approximately the mean thermal speed, the particles of the electron gas respond in such a way that when a steady state is finally established, the field of the particle is screened out within a distance of the order of λ_D. In this way, we obtain a more detailed understanding of how the screening, discussed in Sec. III, is brought about. We then show that if v_0 is more than approximately the mean thermal speed, the field of the particle continues to be screened, but a new phenomenon appears, *viz.*, the excitation of a wake consisting of collective oscillations that carry energy away from the particle. This wake resembles the Čerenkov radiation obtained when fast electrons go through a dielectric. We find that the energy loss to the collective oscillations is of the same order of magnitude as the loss caused by short-range Coulomb collisions with the individual particles.

We begin by Fourier-analyzing the charge density of the specified particle, moving with a constant velocity, $\mathbf{v}_0$

$$-e\rho_s=-e\delta(\mathbf{x}-\mathbf{v}_0t)=-e\textstyle\sum_k e^{ik\cdot(\mathbf{x}-\mathbf{v}_0t)}. \tag{41}$$

Our problem is to calculate the response of the q_k to the field produced by ρ_s. We recall that

$$q_k=(\omega_P^2/2\omega)(\xi_{k,\omega}-\xi_{k,-\omega}). \tag{42}$$

We first find the response of the $\xi_{k,\omega}$ to ρ_s. We have

$$\xi_{k,\omega}=\sum_i\left\{\frac{-i(\mathbf{k}\cdot\mathbf{v}_i)}{\omega-\mathbf{k}\cdot\mathbf{v}_i}+\frac{\mathbf{k}\cdot\dot{\mathbf{v}}_i}{[\omega-(\mathbf{k}\cdot\mathbf{v}_i)]^2}\right\}e^{-ik\cdot\mathbf{x}_i}. \tag{42a}$$

The force on the ith particle arises from two sources: the other particles in the electron gas, and the specified

particle. When the latter is not present, we have, according to (20),

$$\xi_{k,\omega} = -i\omega\xi_{k,\omega}.$$

The external particle changes $\dot{\mathbf{v}}_i$, according to (6) and (41), by

$$\dot{v}_i^{(e)} = -(4\pi e^2/m)i\sum_{k'}[\mathbf{k}'/(k')^2] \times \exp[i\mathbf{k}'\cdot(\mathbf{x}_i - \mathbf{v}_i t)]. \quad (43)$$

Thus we obtain, using the random phase approximation,

$$\dot{\xi}_{k,\omega} = -i\omega\xi_{k,\omega} - (4\pi e^2/m)i\sum_i[1/(\omega - \mathbf{k}\cdot\mathbf{v}_i)^2]e^{-i\mathbf{k}\cdot\mathbf{v}_0 t}. \quad (44)$$

Applying the dispersion relation, (23), we have

$$\dot{\xi}_{k,\omega} = -i\omega\xi_{k,\omega} - ie^{-i\mathbf{k}\cdot\mathbf{v}_0 t}. \quad (45)$$

From (42) we then get

$$\dot{q}_k = -i(\omega_P^2/2)(\xi_{k,\omega} + \xi_{k,-\omega}).$$

To obtain $\ddot{q}_k$, we differentiate again, obtaining

$$\ddot{q}_k = -\omega(\omega^2_P/2)(\xi_{k,\omega} - \xi_{+k,-\omega}) - \omega_P^2 e^{-i\mathbf{k}\cdot\mathbf{v}_0 t},$$

and hence

$$\ddot{q}_k + \omega^2 q_k = -\omega_P^2 e^{-i\mathbf{k}\cdot\mathbf{v}_0 t}. \quad (46)$$

Equation (46) describes a forced harmonic oscillation. A particular solution which describes the steady-state response of the Fourier component of the collective part of the charge density to the field of the specified particle is given by[17]

$$q_k = -[\omega_P^2/\{\omega^2 - (\mathbf{k}\cdot\mathbf{v}_0)^2\}]e^{-i\mathbf{k}\cdot\mathbf{v}_0 t}. \quad (47)$$

We exclude for the time being the case in which the denominator of (47) vanishes. To obtain the Fourier component of the total charge density associated with the specified particle, we must add the density (45) coming from the charge itself. We get

$$\rho_{k_0} = \{1 - [\omega_P^2/\{\omega^2 - (\mathbf{k}\cdot\mathbf{v}_0)^2\}]\}e^{-i\mathbf{k}\cdot\mathbf{v}_0 t}. \quad (48)$$

The charge density as a function of position is

$$\rho_0(\mathbf{x}) = \sum_k \frac{[\omega^2 - \omega_P^2 - (\mathbf{k}\cdot\mathbf{v}_0)^2]}{\omega^2 - (\mathbf{k}\cdot\mathbf{v}_0)^2} e^{i\mathbf{k}\cdot(\mathbf{x} - \mathbf{v}_0 t)}. \quad (49)$$

However, Eq. (49) is formally equivalent to $\eta_e(\mathbf{x})$ (see Eq. (35)). Thus we can conclude that when a particle moves through the electron gas with less than the mean thermal speed, the collective response to its electric field is just such as to screen out that field within a distance of the order of magnitude of λ_D. This is the origin of the screening cloud discussed in the previous section.

[17] In obtaining this solution we have taken $\mathbf{v}_0$ as constant. This will be a good approximation as long as the change in velocity during the period of an oscillation is small compared to $\mathbf{v}_0$, which is the case in most electron gases. See Section V.

We now consider a particle moving with a velocity $\mathbf{v}_0$ such that the denominator of (47) can vanish for some $k < k_D$, i.e., such that

$$(\mathbf{k}\cdot\mathbf{v}_0)^2 = \omega^2 \cong \omega_P^2 + k^2\langle v^2\rangle_{\text{Av}}. \quad (50)$$

For those values of $\mathbf{k}$ such that (50) is satisfied, a steady-state solution for q_k is impossible, since as can be seen from (46), the correct solution corresponds to a resonant excitation of the appropriate q_k. Let us now find the value of $\mathbf{k}$ and $\mathbf{v}_0$ such that (50) can be satisfied. Taking k_x in the direction of $\mathbf{v}_0$, we obtain

$$k_x^2(v_0^2 - \langle v^2\rangle_{\text{Av}}) = \omega_P^2 + (k_x^2 + k_y^2)\langle v^2\rangle_{\text{Av}}. \quad (51)$$

It is immediately evident that real solutions of (51) cannot exist when v_0^2 is less than $\langle v^2\rangle_{\text{Av}}$, so that particles moving with less than the mean thermal speed will not excite collective oscillations. The minimum value of v_0 for which collective oscillations can be excited may be found by setting $k_x = k_y = 0$, and k_z equal to its maximum value k_D. The criterion for excitation becomes

$$v_0^2 \geq (4/3)\langle v^2\rangle_{\text{Av}}, \quad (52)$$

where we have used (15).

We must now solve (46) with the correct boundary conditions for the resonant case. The proper choice of boundary conditions can be seen from the following considerations. The group velocity of the collective oscillations, which measures the speed with which a disturbance is propagated, is given by

$$v_g = (d\omega/dk) \cong [k\langle v^2\rangle_{\text{Av}}/\omega].$$

The maximum value of v_g, obtained by setting $k = k_D$, is $[\frac{3}{4}\langle v^2\rangle_{\text{Av}}]^{\frac{1}{2}}$. A particle moving with a speed high enough to excite oscillation therefore runs ahead of the disturbance that it creates. Hence, as is well known, the disturbance will take the form of a wake trailing behind the particle. The correct boundary condition for excitation is, therefore, that no disturbance exist ahead of the particle. This situation can be contrasted to that existing when $v_0^2 < \frac{3}{4}\langle v^2\rangle_{\text{Av}}$. In the latter case the disturbance can propagate ahead of the particle, so that the field around the particle ultimately reaches a steady state. (Actually a steady state is reached as long as the condition (50) cannot be satisfied, but the use of the expansion (24) for ω^2 is not permissible near the Debye length. This is the origin of the discrepancy between our estimated maximum speed of transmission of a disturbance $[\frac{3}{4}\langle v^2\rangle_{\text{Av}}]^{\frac{1}{2}}$ and minimum speed, $(v_0 = [(4/3)\langle v^2\rangle_{\text{Av}}]^{\frac{1}{2}})$, at which excitation of collective oscillation can occur.)

To solve the boundary condition problem, we first consider $q(\mathbf{x})$ obtained from the q_k given in (46) by

$$q(\mathbf{x}) = \sum_{k<k_D} q_k e^{i\mathbf{k}\cdot\mathbf{x}} = \sum_{k<k_D} \{-\omega_P^2/[\omega^2 - (\mathbf{k}\cdot\mathbf{v}_0)^2]\} e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)}. \quad (52a)$$

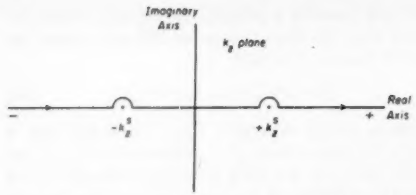


FIG. 1. Contour for the evaluation of the integral in Eq. (54).

When we choose k_z in the direction of $\mathbf{v}_0$, we obtain:

$$q(\mathbf{x})=\sum_{k<k_D}\frac{-\omega_P^2\exp\{i[k_x x+k_y y+k_z(z-v_0 t)]\}}{\omega_P^2+(k_x^2+k_y^2)\langle v^2\rangle_{Av}+k_z^2[\langle v^2\rangle_{Av}-v_0^2]}, \quad (52b)$$

where we have expanded ω^2 according to (24). The denominator in (52b) vanishes for $k_z=k_z^s$, where

$$k_z^s=\pm[(\omega_P^2+(k_x^2+k_y^2)\langle v^2\rangle_{Av})/(v_0^2-\langle v^2\rangle_{Av})]^{\frac{1}{2}}. \quad (53)$$

In the summation over k_z it will be seen to be mathematically convenient to extend our limits from $\pm(k_D^2-k_x^2-k_y^2)^{\frac{1}{2}}$ to $\pm\infty$. This will certainly be justified for v_0 sufficiently larger than $\langle v^2\rangle_{Av}^{\frac{1}{2}}$ and will provide an order of magnitude estimate for all values of v_0. For with v_0 sufficiently large, almost all values of k_z^s will lie within the limits $\pm(k_D^2-k_x^2-k_y^2)^{\frac{1}{2}}$, and the denominator of (52b) will be large for the values of k_z outside these limits. Hence, in this case a negligible error is introduced by extending the range of integration to $\pm\infty$. (For v_0 comparable with $\langle v^2\rangle_{Av}^{\frac{1}{2}}$, the error may become appreciable, and a more detailed calculation is required.) Thus we obtain, on changing our sum to an integral,

$$q(\mathbf{x})=-(1/2\pi)^3\int dk_x\int dk_y\int_{-\infty}^{\infty}dk_z \times\left\{\frac{\omega_P^2\exp[i(k_x x+k_y y+k_z(z-v_0 t)]}{\omega_P^2+(k_x^2+k_y^2)\langle v^2\rangle_{Av}+k_z^2(\langle v^2\rangle_{Av}-v_0^2)}\right\}. \quad (54)$$

The boundary conditions may be introduced, according to well-known techniques, by integrating over the complex k_z plane and choosing a suitable path of integration. It has been shown that $q(\mathbf{x})$ should correspond to a wave moving behind the particle. To insure that there is no disturbance ahead of the particle, we require that $q(\mathbf{x})$ vanish for positive $(z-v_0 t)$. It is readily verified that the proper contour is that given in Fig. 1. For positive $z-v_0 t$ we close the contour by integrating over a large semicircle in the upper half-plane, and thus obtain zero in accordance with our boundary condition. For negative $z-v_0 t$ we close the contour with a large semicircle in the lower half-plane. The contour can be shrunk down to the two singularities, and the residues yield,

$$q(\mathbf{x})=-\frac{\omega_P^2}{4\pi^2[v_0^2-\langle v^2\rangle_{Av}]}\int\int dk_x dk_y \times\left\{\frac{\{\sin[k_z^s(z-v_0 t)]\}e^{i(k_x x+k_y y)}}{\omega_P^2+(k_x^2+k_y^2)\langle v^2\rangle_{Av}}\right\}. \quad (55)$$

The precise form of $q(x)$ is of no great interest to us here, as we shall be mainly interested in computing the energy given up by the incident particle to the wake. This is determined by the reaction of the electric field of the wake on the emitting particle located at $\mathbf{x}=\mathbf{v}_0 t$ or $x=0$, $y=0$, $z=v_0 t$.[18] We need only consider the z component $\mathcal{E}_z$, since by symmetry the x and y components vanish. $\mathcal{E}_z$ vanishes for $z>v_0 t$ and is finite for $z<v_0 t$. There is a discontinuity at $z=v_0 t$, and hence, according to a well-known property of Fourier series, the correct value at $z=v_0 t$ is the sum taken by approaching this point from both sides. Thus we take $(\frac{1}{2})\mathcal{E}_z(v_0 t)$ as approached from the left. A little algebra shows that

$$\mathcal{E}_z(v_0 t)=-\frac{e\omega_P^2}{2\pi}\int\int dk_x dk_y\left\{\frac{1}{\omega_P^2+(k_x^2+k_y^2)v_0^2}\right\}. \quad (56)$$

The force on the particle caused by the wake is thus

$$F_z=-e\mathcal{E}_z(v_0 t). \quad (57)$$

On performing the integration, using polar coordinates, and taking

$$(k_x^2+k_y^2)_{\max}\cong(\tfrac{2}{3})k_D^2=2\omega_P^2/\langle v^2\rangle_{Av},$$

we get:

$$F_z=(e^2\omega_P^2/2v_0^2)\ln(1+2v_0^2/\langle v^2\rangle_{Av}). \quad (58)$$

The limits of integration of (56) were loosely defined, but since the integral has a logarithmic dependence on the actual limits the result is insensitive to the exact limits taken. Thus the rate of energy loss per unit distance dE/dz which is equal to F_z is

$$(dE/dz)^{(1)}=(\pi ne^4/E_0)\ln(1+2v_0^2/\langle v^2\rangle_{Av}), \quad (59a)$$

where E_0 is the energy of the incident particle.

Thus far we have only considered the energy given up by the specified particle to the collective oscillations. It can also transfer energy to the individual particles in short-range collisions. The local electric field around each particle in the electron gas, as we have seen, corresponds to a screened Coulomb potential of range λ_D arising from the η_k contribution to the charge density.

[18] We here are calculating the effects of the $q(\mathbf{x})$ on the particle at a distance for which the concept of a $q(\mathbf{x})$ is clearly not applicable, since we are certainly less than a Debye length removed from the particle. Nevertheless the reacting force may be obtained correctly in this way. For the energy is carried away by the radiation at long distances, where the $q(\mathbf{x})$ can correctly be applied. Because energy is conserved, this energy can come only from the incident particle. Therefore, by formally extrapolating the theory into distances below those for which its physical application is correct, we are able to obtain the correct energy loss.

When the specified particle enters the range of this force it can transfer energy by means of a collision. This transfer is essentially independent of the excitation of collective oscillations, and can best be calculated with aid of the usual collision theory applied to the individual particles. One finds that the energy loss resulting from these short-range collisions is approximately

$$(dE/dz)^{(2)} = (2\pi ne^4/E_0)\ln(\lambda_D E_0/e^2). \quad (59b)$$

It is clear that the two modes of energy loss, described by (59a) and (59b), will generally be of the same order of magnitude.

Experiments by Ruthemann[6] and Lang[7] on energy loss of electrons of energies of the order of 1 to 10 kev in thin films of metal tend to confirm our theoretical prediction of Čerenkov-like radiation. Although the quantum theory should be used in a rigorous treatment of a metal (and will be used in a subsequent paper), the right results for this problem can be obtained by using a correspondence principle argument. We first note that electrons must emit energy in the form of quanta with $E=\hbar\omega$. The probability of emission of a quantum may be estimated from the requirement that the mean energy radiated be equal to the classically calculated value. We then obtain for the number of quanta emitted per unit path length,

$$(dN/dx) = (1/\hbar\omega)(dE/dx).$$

The mean free path for emission of a quantum is the reciprocal of this expression, and is thus given by

$$\lambda = \frac{E_0\hbar\omega}{\pi ne^4\ln(1+2v_0^2/\langle v^2\rangle_{\text{Av}})}.$$

We might expect that the fundamental quantum of energy loss should be quite close to $\hbar\omega_P$. To see this, we note first, using Eq. (53), that for large v_0, k_z must be close to ω_P/v_0, which is small. From Eq. (56) we see that the major contribution to the decelerating field comes when $(k_x^2+k_y^2)\lesssim(\omega_P^2/v_0^2)$. Thus we conclude that $\omega^2=\omega_P^2+k^2\langle v^2\rangle_{\text{Av}}$ must be close to $\omega_P^2[1+(\langle v^2\rangle_{\text{Av}}/v_0^2)]\cong\omega_P^2$. The experiments quoted previously indicate that the electrons actually do lose energy in integral multiples of a fairly sharply defined basic unit. For Al this unit is 14.7 ev, and for Be 19.0 ev. Our calculated $\hbar\omega_P$ for these metals, under the assumption that the valence electrons are all free, is 15.9 ev for Al and 18.8 ev for Be. These results are in remarkably good agreement with experiment, since the effective number of free electrons probably differs somewhat from the number of valence electrons.

Lang has given data on the thickness of his films, from which one can conclude that the mean free path for emission of a quantum is somewhat less than 185A. This compares favorably with our theoretical value of ~150A for this case. A more detailed study of these problems will be given in a subsequent paper.

In conclusion, we now show, without giving a detailed calculation, that the collective response $q(\mathbf{x})$ screens the field of a specified particle of any velocity. To do this, we show that the total responding charge $-eq(\mathbf{x})$ integrated over a small sphere surrounding the incident particle is nearly equal to, and of opposite sign, to that of the incident particle.

From (52a), we obtain for such an integral over a sphere of radius R surrounding $\mathbf{x}=\mathbf{v}_0 t$

$$I = -\int_{\text{sphere}} dx\, q(\mathbf{x})e = \sum_{k<k_D}\left\{\frac{-\omega_P^2 e}{\omega^2-(\mathbf{k}\cdot\mathbf{v}_0)^2}\right\}\int_0^R r^2 dr \times\int_0^\pi d(\cos\vartheta)\int_0^{2\pi} d\varphi e^{ikr\cos\vartheta}. \quad (60)$$

Carrying out the angular integrations, we have

$$I = \sum_{k<k_D} -e\left[\omega_P^2/\{\omega^2-(\mathbf{k}\cdot\mathbf{v}_0)^2\}\right] \times\int_0^R dr 4\pi\{r[\sin(kr)]/k\}, \quad (61)$$

and hence,

$$I = \sum_{k<k_D} e\left[\omega_P^2/\{\omega^2-(\mathbf{k}\cdot\mathbf{v}_0)^2\}\right] \times\{4\pi R^2\{[\cos(kR)]/(kR)\}\}. \quad (62)$$

The expression inside the bracket is a sharply peaked function of k which approaches a Dirac δ function in k space at large R, and is negligible for $kR\gtrsim 1$. For sufficiently large R, the sum over k yields just the value of $\omega_P^2/[\omega^2-(k\cdot v_0)^2]$ at $k=0$, which is unity. Thus for a sufficiently large R, which we denote by R_e, the total charge associated with the collective response is equal in magnitude and opposite in sign to the incident charge. Thus the incident charge is screened out within a distance R_e. The order of magnitude of R_e, the effective screening radius, may be seen from the following argument. In order that the charge be screened, the half-width $k\cong(1/R_e)$ must be sufficiently narrow so that the ratio $\omega_P^2/[\omega^2-(\mathbf{k}\cdot\mathbf{v}_0)^2]$ shall be close to unity. If there is no singularity in the denominator, this ratio will be close to unity for k up to $\sim k_D$, so that $R_e\cong\lambda_D$. If there is a singularity, then the ratio will be unity only up to the neighborhood of $k=k_s$ or k_D, whichever is the smaller. Thus $R_e=\lambda_D$ or $1/k_s$, whichever is the larger. Since high velocity incident particles may have a small k_s, in these cases the screening will not be as good as for a low velocity particle.

We conclude that the field of each particle in the electron gas is screened as the result of the collective response of all the other particles. For most particles, the field is effectively screened within a distance of approximately λ_D. For those few particles with much higher than mean thermal speed, the screening radius may be somewhat greater than λ_D. These particles also excite collective oscillation, and in fact this excitation

is one of the means of bringing such particles into thermal equilibrium with the rest of the assembly. The excitation of collective oscillations is a consequence of the long-range part of the Coulomb interactions; one may regard it as a description of the effect of the simultaneous many-body collisions brought by the long range of the force. The remaining screened short-range Coulomb interactions can, as we have seen, be described in terms of the usual two-body collision theory. Thus we see that the use of an individual particle model for an electron gas is justified in the treatment of phenomena involving distances less than the Debye length.

V. THE ROLE OF INDIVIDUAL PARTICLES IN ORGANIZED BEHAVIOR

In this section we wish to develop a physical picture of the motion of the individual particles as they take part in the collective motion. We first consider the state of the electron gas in the absence of collective oscillation, i.e., when the q_k and $\dot{q}_k$ are zero for all k. In this state, we have

$$\sum_i [\omega_P^2/\{\omega^2-(\mathbf{k}\cdot\mathbf{v}_{0i})^2\}]e^{-i\mathbf{k}\cdot\mathbf{x}_{0i}}=0, \quad (63a)$$

$$\sum_i [(\mathbf{k}\cdot\mathbf{v}_{0i})\omega_P^2/\{\omega^2-(\mathbf{k}\cdot\mathbf{v}_{0i})^2\}]e^{-i\mathbf{k}\cdot\mathbf{x}_{0i}}=0, \quad (63b)$$

where $\mathbf{x}_{0i}$ and $\mathbf{v}_{0i}$ are the position and velocity of the ith particle in the absence of organized oscillation. The state of no collective oscillation described by (63a) and (63b) differs from a state in which the particles move in a completely random way, as would exist in the absence of interaction.

In the absence of interaction the particles move in straight lines with constant velocities. This motion produces random fluctuations in the particle density which make the most likely absolute value of q_k different from zero. To see this, consider $\langle|q_k|^2\rangle_{\text{Av}}$ which is, according to (31),

$$\langle|q_k|^2\rangle_{\text{Av}}=\left\langle\sum_{ij}\frac{\omega_P^4\exp[i\mathbf{k}\cdot(\mathbf{x}_i-\mathbf{x}_j)]}{[\omega^2-(\mathbf{k}\cdot\mathbf{v}_i)^2][\omega^2-(\mathbf{k}\cdot\mathbf{v}_j)^2]}\right\rangle_{\text{Av}}. \quad (64)$$

Since the particles are distributed at random, only those terms with $i=j$ contribute to the average, and for small k one then obtains approximately.

$$\langle|q_k|^2\rangle_{\text{Av}}=n. \quad (65)$$

Thus we see that the particles in the state of lowest collective energy cannot move in straight lines and possess a random distribution of positions. Instead there must be fluctuations in particle velocities resulting from the Coulomb interaction which lead to just such correlations of particles positions as to make $q_k=0$. For this state, the charge density ρ_k reduces to η_k, the individual particles part, which for small k is much less than ρ_k. Thus these correlations produce the screening of the field of each individual particle by the electron gas.

Since for small k, q_k is very nearly equal to ρ_k, we see that the state with $q_k=0$ corresponds to one in which the density fluctuations are greatly reduced below the values to be expected in a random distribution. The state $q_k=0$ corresponds to the state of lowest possible collective energy, but does not correspond to a state of thermal equilibrium. It will be shown in Appendix I, however, that the state of thermal equilibrium is not far from the state of lowest collective energy. There we evaluate the probability of fluctuation in the charge density for a state of thermal equilibrium, and show that as a consequence of the Coulomb repulsion, this probability is greatly reduced below that present in the absence of interaction. Thus we verify in an independent way the reduction in density fluctuations implied by $q_k=0$.

It is difficult to solve in detail for the motion of the particle in the state of zero oscillation. One of the advantages of the collective description is that detailed solutions for this motion are not required. Since each particle moves in the screened Coulomb field of the other particles, its motion will usually not differ markedly from straight line motion. An estimate of the validity of the uniform straight line motion assumption may be obtained by comparing the mean-free path for collision (because of the screened Coulomb force) with the interparticle distance. It should be emphasized here that the assumption that the particles in an electron gas are approximately independent (free) is valid only because the organized behavior acts to screen out the long-range part of the Coulomb interparticle force.

Let us now consider the way in which the motion of the individual particles changes as a result of collective oscillation. We solve for the motion of the particle under the assumption that its motion in the absence of oscillation can be approximated as uniform, rectilinear motion. This approximation will be suitable when the fractional fluctuations of the particle velocity resulting from the short-range screened Coulomb interparticle forces are small during the period of an oscillation. These fluctuations will be small provided the mean-free path for short-range collisions is considerably greater than the distance covered by the particle during this period, which is approximately λ_D. This criterion is satisfied in all electron gases of interest. Under these circumstances the short-range collisions give rise only to a small damping of the organized oscillation, with a damping time $\tau=\tau_0(l/\lambda_D)$, where τ_0 is the period of an oscillation and l is the mean free path.[19] When l is of the order of magnitude of λ_D, the whole concept of organized oscillations becomes doubtful.

In the straight-line approximation we have, in the absence of oscillation,

$$\dot{\mathbf{x}}_{0i}=\mathbf{v}_{0i}, \quad \dot{\mathbf{v}}_{0i}=0. \quad (66)$$

For small amplitudes of collective oscillation, there is a correspondingly small disturbance of the motion of each

[19] For a discussion of damping, see D. Bohm and E. P. Gross, Paper B, reference 5.

particle. Thus we may write

$$\mathbf{x}_i=\mathbf{x}_{0i}+\delta\mathbf{x}_i, \quad \mathbf{v}_i=\mathbf{v}_{0i}+\delta\mathbf{v}_i. \tag{67}$$

We obtain the equations of motion for $\delta\mathbf{x}_i$ and $\delta\mathbf{v}_i$ by substituting (67) into the equations of motion (6), using $\rho_k=q_k+\eta_k$. However, the use of the straight-line approximation for the motion of the particles in the absence of oscillation is equivalent to neglecting the force arising from the individual particles part of the charge density, η_k. Thus we may write

$$\delta\dot{\mathbf{x}}_i=\delta\mathbf{v}_i, \tag{68}$$

$$\delta\dot{\mathbf{v}}_i=-(4\pi e^2/m)i\textstyle\sum_k(\mathbf{k}/k^2)q_k e^{i\mathbf{k}\cdot(\mathbf{x}_{0i}+\delta\mathbf{x}_i)}. \tag{69}$$

We may expand the exponential in (69), and, in the linear approximation, neglect terms like $q_k(\mathbf{k}\cdot\delta\mathbf{x}_i)$ and $(\mathbf{k}\cdot\delta\mathbf{x}_i)^2$. Thus (69) becomes

$$\delta\dot{\mathbf{v}}_i=-(4\pi e^2/m)i\textstyle\sum_k(\mathbf{k}/k^2)q_k e^{i\mathbf{k}\cdot\mathbf{x}_{0i}}. \tag{70}$$

Using $q_k=(\omega_P^2/2\omega)(\xi_{k,\omega}-\xi_{k,-\omega})$ one can then easily verify that the following represent an approximate steady-state solution of (68) and (69),

$$\delta\mathbf{x}_i\cong i\frac{2\pi e^2\omega_P^2}{m}\sum_k\left\{\frac{\xi_{k,\omega}}{(\omega-\mathbf{k}\cdot\mathbf{v}_{0i})^2}-\frac{\xi_{k,-\omega}}{(\omega+\mathbf{k}\cdot\mathbf{v}_{0i})^2}\right\}\frac{\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{x}_{0i}}}{k^2\omega}. \tag{71}$$

$$\delta\mathbf{v}_i=\frac{2\pi e^2\omega_P^2}{m\omega}\sum_k\frac{\mathbf{k}}{\omega k^2}\left\{\frac{\xi_{k,\omega}}{\omega-\mathbf{k}\cdot\mathbf{v}_{0i}}+\frac{\xi_{k,-\omega}}{\omega+\mathbf{k}\cdot\mathbf{v}_{0i}}\right\}e^{i\mathbf{k}\cdot\mathbf{x}_{0i}}. \tag{72}$$

From (71) and (72), we see that in a state of collective oscillation there is a small wave-like perturbation in the position and velocity of each particle, possessing a definite phase relation to the collective coordinates $\xi_{k,\omega}$ and $\xi_{k,-\omega}$ (i.e., q_k and $\dot{q}_k$). We may calculate the additional contribution to the charge density resulting from this perturbation. This is

$$\delta\rho_k=\textstyle\sum_i\{e^{-i\mathbf{k}\cdot(\mathbf{x}_{0i}+\delta\mathbf{x}_i)}-e^{-i\mathbf{k}\cdot\mathbf{x}_{0i}}\}$$

$$\cong\textstyle\sum_i e^{-i\mathbf{k}\cdot\mathbf{x}_{0i}}(-i\mathbf{k}\cdot\delta\mathbf{x}_i) \tag{73}$$

in the linear approximation. Substituting for $\delta\mathbf{x}_i$ from (71) and using the random phase approximation, we obtain

$$\delta\rho_k\cong\frac{2\pi e^2\omega_P^2}{m\omega}\sum_i\left\{\frac{\xi_{k,\omega}}{(\omega-\mathbf{k}\cdot\mathbf{v}_{0i})^2}-\frac{\xi_{k,-\omega}}{(\omega+\mathbf{k}\cdot\mathbf{v}_{0i})^2}\right\}. \tag{74}$$

In order that the part of the charge density arising from the response of the individual particles to the collective part of the charge density q_k should be equal to this same collective charge density, it is necessary that the dispersion relation,

$$1=(4\pi e^2/m)\textstyle\sum_i\{1/(\omega-\mathbf{k}\cdot\mathbf{v}_{0i})^2\}, \tag{75}$$

be satisfied for both $\pm\omega$, as may be verified using (42). But within the linear approximation the dispersion relation (75) is identical with the dispersion relation (23) derived earlier. Thus our dispersion relation guarantees the self-consistency of our assumption of the collective oscillation arising from the cumulative small response of the individual particles to the collective field.

The consistency of the separation of the charge density into q_k and η_k further requires that η_k, in the linear approximation, be independent of q_k. Thus the change in η_k resulting from the $\delta\mathbf{x}_i$ and the $\delta\mathbf{v}_i$ should vanish. This change is, by our definition of η_k, (30),

$$\delta\eta_k=\sum_i\left[\left\{1-\frac{\omega_P^2}{\omega^2-(\mathbf{k}\cdot\mathbf{v}_{0i})^2}\right\}\{-i\mathbf{k}\cdot\delta\mathbf{x}_i\}\right.$$

$$\left.-\left\{\frac{2\omega_P^2\mathbf{k}\cdot\mathbf{v}_{0i}\mathbf{k}\cdot\delta\mathbf{v}_i}{[\omega^2-(\mathbf{k}\cdot\mathbf{v}_{0i})^2]^2}\right\}\right]e^{-i\mathbf{k}\cdot\mathbf{x}_{0i}}. \tag{76}$$

On substituting for $\delta\mathbf{x}_i$ and $\delta\mathbf{v}_i$ from (71) and (72), and using the random phase approximation it is readily established that $\delta\eta_k=0$.

VI. RANDOM PHASE APPROXIMATION

In this section we justify our use of the random phase approximation, which is perhaps the central approximation in our treatment of the electron gas. We justify this approximation in detail for a specific case. The justification of the approximation as used elsewhere in the paper can be carried out using similar methods.

We first encounter the random phase approximation in Eq. (8), where we neglect the term

$$R=\sum_{\substack{ijk'\\k'\neq k}}\frac{4\pi e^2}{m}\frac{\mathbf{k}\cdot\mathbf{k}'}{(k')^2}\exp[i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{x}_i]\exp[-i\mathbf{k}'\cdot\mathbf{x}_j]$$

$$=\frac{4\pi e^2}{m}\sum_{\substack{k'\\k'\neq k}}\frac{\mathbf{k}\cdot\mathbf{k}'}{(k')^2}\rho_{k'-k}\rho_{k'}, \tag{77}$$

in comparison with the term

$$S=(4\pi e^2/m)\textstyle\sum_j e^{-i\mathbf{k}\cdot\mathbf{x}_j}. \tag{78}$$

If the particles were distributed at random, the average of R would be zero. However, the value of R would tend to undergo random fluctuations as a result of the random motions of the particles. These fluctuations would produce a corresponding random wavering of the frequency of plasma oscillations, and would lead to coupling between oscillations of different wavelengths, (as can be seen from (77) and (8)). The random phase approximation will be justified provided we can show that the effects of these fluctuations can be neglected.

Before we attempt to estimate the size of R, it will be convenient to rewrite it in a form which groups together certain pairs of terms having definite phase relations. To do this, we interchange i and j in (77) and replace $\mathbf{k}'$ by $\mathbf{k}'-\mathbf{k}$. Since the substitutions do not change the value of the sum, we can re-express R as the

mean of the two expressions. We obtain

$$R=(2\pi e^2/m)\sum_{\substack{ijk'\\ k'\neq k}}\left\{\frac{k^2}{(\mathbf{k}-\mathbf{k}')^2}+\frac{\mathbf{k}\cdot\mathbf{k}'(k^2-2\mathbf{k}\cdot\mathbf{k}')}{(k')^2(\mathbf{k}-\mathbf{k}')^2}\right\}$$
$$\times\exp[i(\mathbf{k}'-\mathbf{k})-(\mathbf{x}_i-\mathbf{x}_j)]\exp(-i\mathbf{k}\cdot\mathbf{k}_j). \quad (79)$$

Let us begin with first term in the above expression, which we rewrite with the aid of the substitution, $\mathbf{k}'-\mathbf{k}=\mathbf{k}''$, as

$$R'=(2\pi e^2k^2/m)\sum_{\substack{ijk''\\ k''\neq 0}}[1/(k'')^2]$$
$$\times\exp[i\mathbf{k}''\cdot(\mathbf{x}_i-\mathbf{x}_j)]\exp(-i\mathbf{k}\cdot\mathbf{x}_j). \quad (80)$$

The expression

$$\sum_{\substack{ijk''\\ k''\neq 0}}[1/(k'')^2]\exp[i\mathbf{k}''\cdot(\mathbf{x}_i-\mathbf{x}_j)]$$

is proportional to the potential of the jth electron in the field of all of the other electrons plus that of the "smeared out" distribution of positive charge. This potential is

$$\varphi(\mathbf{x}_j)=4\pi e\sum_{k\neq 0}(\rho_k/k^2)e^{i\mathbf{k}\cdot\mathbf{x}_j}. \quad (81)$$

Now we wish to calculate the mean fluctuation in this quantity. To do this, we first obtain $[\varphi(\mathbf{x}_j)]^2$, and then average over-all values of ρ_k. We have, from (81),

$$[\varphi(\mathbf{x}_j)]^2=(4\pi e)^2\sum_{kk'}[\rho_k\rho_{k'}/k^2(k')^2]\exp[i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}_j].$$

We shall average ρ_k under the assumption that the system is in thermodynamic equilibrium. This is close to the situation, usually present in practice, in which there are organized oscillations of small amplitude superposed on a background of random thermal motion.[20] From Appendix I [Eqs. (A10) and (A11)] we find that the probability of a given set of ρ_k is

$$W(\cdots\rho_k\cdots)\sim\exp\{-\textstyle\sum_k|\rho_k|^2[(4\pi e^2/k^2\kappa T)+(1/n)]\}$$
$$\times\{\cdots d(r_k{}^2/1)\cdots\}\{\cdots d\varphi_k\cdots\},$$

where $\rho_k=r_ke^{i\varphi_k}$. As shown in Appendix I, this expression includes the effects of the correlations in particle positions brought about by the Coulomb force. Since W does not depend on φ_k products like $\rho_k\rho_{k'}=r_kr_{k'}e^{i(\varphi_k+\varphi_{k'})}$ average out to zero unless $k=-k'$. Using this fact, we find

$$\langle[\varphi(\mathbf{x}_j)]^2\rangle_{Av}=(4\pi e)^2\sum_{k\neq 0}\langle r_k{}^2\rangle_{Av}/k^4. \quad (82)$$

On substituting for $\langle r_k{}^2\rangle_{Av}$ its value $nk^2/[k^2+(4\pi ne^2/\kappa T)]$ as obtained in (A12), we obtain

$$\langle[\varphi(\mathbf{x}_j)]^2\rangle_{Av}=(4\pi e)^2\sum_{k\neq 0}\frac{n}{k^2[k^2+(4\pi e^2/\kappa T)n]}. \quad (83)$$

We transform the sum over k to an integral, and obtain, after integration,

$$\langle[\varphi(\mathbf{x}_j)]^2\rangle_{Av}=(4\pi e)^2(\lambda_D/4\pi)n. \quad (84)$$

From (80) and (84) we see that the fluctuation in R', $\delta R'$ is thus given by

$$\delta R'=(\pi n\lambda_D)^{\frac{1}{2}}e^2\rho_k(k^2/m). \quad (85)$$

This must be compared with the term S given in (78). The ratio is

$$\frac{\delta R'}{S}=\frac{\delta R'}{\omega_P{}^2\rho_k}=\frac{k^2a^{\frac{1}{2}}\lambda_D{}^{\frac{1}{2}}}{4\pi^{\frac{1}{2}}}=\frac{k^2\langle v^2\rangle_{Av}}{\omega_P{}^2}\left(\frac{a}{\lambda_D}\right)^{\frac{1}{2}}\frac{1}{12\pi^{\frac{1}{2}}}, \quad (86)$$

where $n=1/a^3$.

We see that the modification in the equation of motion (8) resulting from the fluctuation in R' is of order $[k^2\langle v^2\rangle_{Av}/\omega_P{}^2]$, but is multiplied by a term of order $\sim(1/20)(a/\lambda_D)^{\frac{1}{2}}$, which is very small in most electron gases.

For a typical density of 10^{12} per cm^3, we have $\lambda_D\sim 10^{-3}$ cm and $a=10^{-4}$ cm. Thus not only does the R' term introduce a very small correction to the net frequency, but this correction is much less than the $[k^2\langle v^2\rangle_{Av}/\omega_P{}^2]$ correction which arises when we consider the effect of thermal motions. In the case described, the fluctuations produce a fractional correction of about 10^{-3} in the $[k^2\langle v^2\rangle_{Av}/\omega_P{}^2]$ term.

The correction $(1/20)(a/\lambda_D)^{\frac{1}{2}}$ would only become appreciable for $\lambda_D\ll a$. Such cases occur in practice only for degenerate electron gases. In these cases, however, Fermi statistics reduce the effects of the fluctuations relative to the Boltzmann estimate, so that the random phase approximation will still be justified. This problem will be considered in our paper extending these results to the quantum theory.

It can be shown by arguments similar to the foregoing that the remaining terms in R are of the same order of magnitude as R' or smaller. A similar justification can be applied for the use of the random phase approximation elsewhere in this paper.

VII. CONCLUSION

In conclusion we give a brief summary of our results in terms of a physical picture of the behavior of the electron gas. As we have seen, the density fluctuations can be split into two approximately independent components, associated, respectively, with the collective and individual particle aspects of the assembly. The collective component, which is present only for wavelengths $>\lambda_D$, represents organized oscillation brought about by the long-range part of the Coulomb interaction. When

[20] The assumption of thermal equilibrium in the estimation of $[\varphi(\mathbf{x}_j)]^2$ necessarily neglects those irreversible effects associated with the damping of the collective oscillation. Actually, the effects of damping must be contained in terms like R, since these effects are not included in the terms we retain in the random phase approximation. However, as was pointed out in Sec. V, for electron gases of interest this damping will be small.

such an oscillation is excited, each individual particle suffers a small perturbation of its velocity and position arising from the combined potential of all the other particles. The contribution to the density fluctuations resulting from these perturbations is in phase with the potential producing it, so that in an oscillation we find a small organized wave-like perturbation superposed on the much larger random thermal motion of the particle. The cumulative potential of all the particles may, however, be considerable because the long range of the force permits a very large number of particles to contribute to the potential at a given point.

The individual particles component of the density fluctuation is associated with the random thermal motion of the particles and shows no collective behavior. It represents the individual particles surrounded by comoving clouds which screen their fields within a distance $\sim\lambda_D$. Thus it describes an assembly of effectively free particles interacting only through the short-range part of the Coulomb force. The screening of the field of a given particle is actually brought about by the Coulomb repulsion which leads to a deficiency of electrons in the immediate neighborhood of the particle. This same process also leads to a large reduction in the random fluctuations of the density in the electron gas for wavelength larger than λ_D.

When we fix our attention on a specific individual electron in the assembly, we may study its interaction with the other electrons by applying our split-up of the density fluctuations to the remainder of the gas. The response of the collective part of the density fluctuation to the field of the specified particle leads to two significant effects: the screening of this particle's field within a distance $\sim\lambda_D$, and the excitation of collective oscillations when the speed of the particle is greater than mean thermal speed. The particle also interacts with the individual particles component of the density fluctuations. However, the field that is the result of this component may be viewed as the sum of the screened fields of individual particles. Thus the interaction of our specified particle with the individual particles component of the density fluctuation can be described in terms of short-range collisions between pairs of particles.

We have used the random phase approximation throughout this paper. The use of this approximation is equivalent to the neglect of the damping and wavering-in-frequency of the collective oscillation resulting from the individual particles character of the electron gas. When the random phase approximation is justified, the collective component and the individual particles component of the density fluctuations will not be significantly coupled, and thus can be treated independently.

The authors wish to thank Dr. Conyers Herring for informing us of the experiments of Ruthemann and Lang. One of us (D.P) would like to acknowledge the partial support of the ONR during the writing of this paper.

APPENDIX I. STATISTICAL MECHANICAL TREATMENT OF DENSITY FLUCTUATIONS IN THE ELECTRON GAS

In this appendix, we shall calculate the mean amplitude of fluctuation of the Fourier components, ρ_k, of the density in an electron gas. The method used here provides an alternative approach to the screening problem, and yields certain results that we apply in the verification of the random phase approximation.

We begin with the well-known statistical mechanical expression for the probability that the electron gas is in a "microscopic" state in which the nth particle lies between $\mathbf{x}_n$ and $\mathbf{x}_n+d\mathbf{x}_n$, while its momentum lies between $\mathbf{p}_n$ and $\mathbf{p}_n+d\mathbf{p}_n$,

$$W'(\mathbf{x}_1, \mathbf{x}_2\cdots\mathbf{x}_n\cdots\mathbf{p}_1\cdots\mathbf{p}_n\cdots) \sim e^{-(E/\kappa T)}d\mathbf{x}_1d\mathbf{x}_2\cdots d\mathbf{x}_n\cdots d\mathbf{p}_1\cdots d\mathbf{p}_n\cdots, \quad \text{(A1)}$$

where E is the total energy of the system, kinetic plus potential,

$$E=\sum_n(p_n^2/2m)+\tfrac{1}{2}\sum_{\substack{m,n\\ m\neq n}} V(\mathbf{x}_m-\mathbf{x}_n).$$

We can carry out the integral over momentum space, obtaining

$$W(\mathbf{x}_1, \mathbf{x}_2\cdots\mathbf{x}_n) \sim\exp\{-\tfrac{1}{2}[\sum_{\substack{m,n\\ m\neq n}} V(\mathbf{x}_m-\mathbf{x}_n)]/\kappa T\}d\mathbf{x}_1\cdots d\mathbf{x}_n\cdots. \quad \text{(A2)}$$

Let us now divide the available space into cells of volume ϵ so small that no important physical property changes within them, but large enough to contain enough particles so that the number of particles in the cell can be regarded as approximately continuous. The smallest suitable value of ϵ corresponds to about 4 or 5 inter-atomic spacings. The state of the system can then be specified in terms of the number of particles $(N_1, N_2\cdots N_i\cdots)$ in each cell. It is readily shown by methods well-known in statistical mechanics that

$$W(N_1\cdots N_i\cdots) \sim\frac{[\exp\{-(1/2\kappa T)\sum_{i\neq j}N_iN_jV(\mathbf{x}_i-\mathbf{x}_j)\}]N!}{(N_1)!(N_2)!\cdots(N_i)!\cdots}, \quad \text{(A3)}$$

where $N=\sum_i N_i$. Here, $\mathbf{x}_i$ refers to the mean coordinate of the ith cell. For the Coulomb potential, we can, as is usually done in electrostatics, neglect the interaction of particles in a given cell. Using Stirling's approximation for $N_i!$ we obtain

$$W(N_1\cdots N_i\cdots)\sim\exp\{N\ln N-\sum_i N_i\ln N_i -(1/2\kappa T)\sum_{i\neq j}N_iN_jV(\mathbf{x}_i-\mathbf{x}_j)\}. \quad \text{(A4)}$$

Let N_0 be the mean number of electrons in a cell. We shall be interested in the small fluctuations, δN_i about

this mean number. Thus we can write

$$N_i = N_0 + \delta N_i. \tag{A5}$$

We now substitute (A5) into (A4) and expand the argument of the exponential in a series of powers of $\delta N/N$, retaining only second-order terms. In doing this, we note that since the electrons are embedded in a uniform distribution of positive charge of density N_0/ϵ, the electrostatic energy is

$$e^2 \sum_{i \neq j} (N_i - N_0)(N_j - N_0)/|\mathbf{x}_i - \mathbf{x}_j| = e^2 \sum_{i \neq j} \delta N_i \delta N_j/|\mathbf{x}_i - \mathbf{x}_j|.$$

We then obtain (using $\sum_i \delta N_i = 0$ and leaving out constant factors)

$$W(N_1 \cdots N_i \cdots) \sim \exp\{-(e^2/2\kappa T) \sum_{i \neq j} \delta N_i \delta N_j/|\mathbf{x}_i - \mathbf{x}_j| - \sum_i (\delta N_i)^2/2N_0\}. \tag{A6}$$

We now expand δN_i as a Fourier series with periodic boundary conditions in a box of unit volume.

$$\delta N_i = \sum_{k \neq 0} \rho_k \epsilon e^{i\mathbf{k}\cdot\mathbf{x}_i}, \tag{A7}$$

with $\rho_k^* = \rho_{-k}$. Since ρ_k is complex, it will be convenient to write it as

$$\rho_k = r_k e^{i\varphi_k}, \tag{A8}$$

with $r_k = r_{-k}$, $\varphi_k = -\varphi_{-k}$. Only half of the ρ_k are independent. It is readily verified that we obtain

$$(e^2/2) \sum_{i \neq j} [\delta N_i \delta N_j/|\mathbf{x}_i - \mathbf{x}_j|] = 2\pi e^2 \sum_k (1/k^2) |\rho_k|^2 = 2\pi e^2 \sum_k (1/k^2) r_k^2, \tag{A9}$$

and

$$W \sim \exp\{-\sum_k [(2\pi e^2/k^2 \kappa T) + (1/2n)] r_k^2\}, \tag{A10}$$

where n is the average electron density.

We obtain the volume element in the space of r_k and φ_k by regarding the N_i as continuous variables, and by assuming that we have chosen $dN_i = 1$ in Eq. (A6). (A6) may therefore be multiplied by the volume element in N_i space $(dN_1, dN_2 \cdots dN_i \cdots)$, which is also unity. The volume element in the space of r_k and φ_k is then given by the Jacobian of the transformation from the N_i space. This is

$$J \sim [d(r_{k1}^2/2) d(r_{k2}^2/2) \cdots d(r_{ki}^2/2)] \times [d\varphi_{k1} \cdots d\varphi_{ki} \cdots]. \tag{A11}$$

The probability WJ is a product of separate functions, one for each k. This indicates that in the approximations used, the statistical fluctuations associated with each k are independent.

We now calculate the mean value of $|\rho_k|^2 = (r_k)^2$. In obtaining this, we note that $r_{-k}^2 = r_k^2$, so that we must replace r_k^2 in (A10) by $2r_k^2$, but sum only over half the values of k. We get

$$\langle r_k^2 \rangle_{Av} = \frac{\int\int \cdots \int W r_k^2 (\cdots d(r_k^2) \cdots)(\cdots d\varphi \cdots)}{\int\int \cdots \int W (\cdots d(r_k^2) \cdots)(\cdots d\varphi_k \cdots)} = \frac{nk^2}{(4\pi ne^2/\kappa T) + k^2}. \tag{A12}$$

For $k^2 \gg k_D^2$ we see that $\langle r_k^2 \rangle_{Av}$ reduces to n, the value which would exist in a random distribution. For $k^2 \ll k_D^2$ we obtain $\langle r_k^2 \rangle_{Av} = nk^2 \lambda_D^2$. Thus, we see that for long distances, the Coulomb forces cause the fluctuations to be greatly reduced.

We can now write down the mean Coulomb potential energy associated with each k, noting from Eq. (A9), that the contributions of different k are independent. This mean potential is

$$\langle V_k \rangle_{Av} = \langle 2\pi e^2 \rho_k^2/k^2 \rangle_{Av} = \tfrac{1}{2}(\kappa T/1 + k^2 \lambda_D^2). \tag{A13}$$

For $k \ll k_D$ the mean potential energy is just $\kappa T/2$. If we recall that for long wavelengths, the ρ_k act almost entirely collectively and undergo simple harmonic motion, then the previous result can easily be understood in terms of the equipartition theorem which states that the mean potential energy of a harmonic oscillator is $\kappa T/2$. For $k \gg k_D$, $\langle V_k \rangle_{Av}$ becomes $2\pi ne^2/k^2$. This is just what is obtained by assuming that the particles move with a random distribution in space. These results show in another application that for short distances, the individual particles point of view gives the right results while for long distances, the organization resulting from the Coulomb forces must be taken into account.

In thermodynamic equilibrium, the energy is actually distributed partly in the collective oscillation and partly in the random thermal motion of the individual particles. Each of the collective modes will have a mean energy of κT. This energy is usually very small in comparison to the collective energy that would be present if one of the modes were excited systematically. For even though each particle makes only a small contribution to the collective modes, the cumulative contributions of all the particles to a given mode may result in energies very much greater than κT. It also follows that the collective coordinate q_k, is not quite equal to zero in the state of thermodynamic equilibrium. But since κT is very much less than what would be present if a given q_k were systematically excited, the state, $q_K = 0$, is not far from the state of thermodynamic equilibrium.

Finally, let us note that the approximation of regarding N_i as continuous can be justified rigorously only for values of k appreciably smaller than $1/a$, where a is the interparticle spacing. Nevertheless, it turns out that for large values of k, this procedure still gives what we know to be the right value of $\langle r_k^2 \rangle_{Av}$, *viz.*, n. This

indicates that (A12) is essentially correct over all values of k.

APPENDIX II. COLLECTIVE DESCRIPTION OF PARTICLE INTERACTIONS

The approach used in Secs. II and III to study the organization brought about by the Coulomb interactions can be extended to a general repulsive particles interaction. In this appendix we indicate briefly the appropriate generalization of the results obtained in the above sections. We assume the interparticle potential is $V(|\mathbf{x}_i-\mathbf{x}_j|)$ and express it as a Fourier series in a box of unit volume with periodic boundary conditions as:

$$V(|\mathbf{x}_i-\mathbf{x}_j|)=\sum_k V_k e^{i\mathbf{k}\cdot(\mathbf{x}_i-\mathbf{x}_j)}. \quad \text{(A14)}$$

The equation of motion of the ith particle is thus

$$\ddot{\mathbf{x}}_i=-(i/m)\sum_{kj}\mathbf{k}V_k e^{i\mathbf{k}\cdot(\mathbf{x}_i-\mathbf{x}_j)}. \quad \text{(A15)}$$

The time variation of the kth density fluctuation is given by

$$d^2\rho_k/dt^2=-\sum_i(\mathbf{k}\cdot\mathbf{v}_i)^2e^{-i\mathbf{k}\cdot\mathbf{x}_i} -\sum_{k'ij}(V_{k'}/m)\mathbf{k}\cdot\mathbf{k}'\{\exp[i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{x}_i]\} \times\{\exp(-i\mathbf{k}'\cdot\mathbf{x}_j)\}. \quad \text{(A16)}$$

We now assume that the random phase approximation may be applied to this case. The use of this approximation cannot be justified in general, but must be considered separately for each law of force. When it is valid Eq. (A16) may be reduced to

$$d^2\rho_k/dt^2=-\sum_i(\mathbf{k}\cdot\mathbf{v}_i)^2e^{-i\mathbf{k}\cdot\mathbf{x}_i} -\sum_j(nV_k/m)k^2e^{-i\mathbf{k}\cdot\mathbf{x}_j}. \quad \text{(A17)}$$

The physical significance of the terms on the right-hand side of (A17) is the same as for the corresponding terms in Eq. (9): the first term represents the contributions from the random thermal motions of individual particles; the second term represents the effects of particles interaction and can lead to organized behavior.

Under conditions such that the individual particles term can be neglected in (A17), we have

$$d^2\rho_k/dt^2+(nk^2V_k/m)\rho_k=0, \quad \text{(A18)}$$

and the ρ_k carry out oscillations of frequency,

$$\omega=(nk^2V_k/m)^{\frac{1}{2}}. \quad \text{(A19)}$$

The criterion that the ρ_k display predominantly collective behavior, and hence that a collective description be appropriate, is

$$(nk^2V_k/m)\gg\langle(\mathbf{k}\cdot\mathbf{v}_i)^2\rangle_{\mathrm{Av}}. \quad \text{(A20)}$$

For an isotropic velocity distribution, Eq. (A20) becomes

$$nV_k\gg(2/3)\mathcal{T}, \quad \text{(A21)}$$

where $\mathcal{T}$ denotes the mean particle kinetic energy. Thus we see that a strong interaction and a high particle density favor organized behavior, while high random thermal velocities oppose it.

Just as was done for the Coulomb interactions, we can find a collective coordinate q_k which oscillates harmonically when random thermal motions are considered, and thus obtain a more exact dispersion relation than Eq. (A19). The appropriate normalized, collective coordinate is

$$q_k=\sum_i\frac{(nk^2V_k/m)}{\omega^2-(\mathbf{k}\cdot\mathbf{v}_i)^2}e^{-i\mathbf{k}\cdot\mathbf{x}_i}, \quad \text{(A22)}$$

where ω satisfies the following dispersion relation

$$1=\sum_j(k^2V_k/m)/[\omega-(\mathbf{k}\cdot\mathbf{v}_j)]^2, \quad \text{(A23)}$$

We may obtain an approximate solution for (A23) in the limit of small $(\mathbf{k}\cdot\mathbf{v}_j/\omega)$ which is essentially the same limit as that in which the criterion (A21) is satisfied. The expansion of (A23) in powers of $(\mathbf{k}\cdot\mathbf{v}_j/\omega)$ yields

$$\omega^2=(nk^2V_k/m)+(m\langle v^2\rangle_{\mathrm{Av}}/nV_k), \quad \text{(A24)}$$

which for sufficiently small $(\mathcal{T}/nV_k)$ reduces to the dispersion relation obtained earlier (Eq. (A19)).

It is also possible to split up the density fluctuation into its collective component q_k, and an individual particles component η_k. It can easily be verified that with the choice (A22) for q_k, one obtains:

$$\rho_k=q_k+\eta_k, \quad \text{(A25)}$$

with

$$\eta_k=\sum_i\frac{\omega^2-(\mathbf{k}\cdot\mathbf{v}_i)^2-(k^2V_kn/m)}{\omega^2-(\mathbf{k}\cdot\mathbf{v}_i)^2}e^{-i\mathbf{k}\cdot\mathbf{x}_i}. \quad \text{(A26)}$$

It may be seen that when the criterion for collective behavior (A21) is satisfied, the corresponding density fluctuation is essentially collective, and may be described by q_k.

Theory of Conductivity of Semiconductors*

GEORGE JAFFÉ
Louisiana State University, Baton Rouge, Louisiana

(Received February 26, 1951)

The author's older theory of conduction in polarizable media is generalized in such a way as to include rectification. An electronic component (positive holes or electrons) and an ionic component (negative or positive ions) are taken into account. The former are subjected to the boundary conditions of the Mott-Schottky theory, the latter to the boundary conditions of the author's theory of polarization. Significant deviations from the Mott-Schottky theory are caused by the ionic component.

In Sec. I the fundamental equations are established in a general way and the boundary conditions are discussed. In the remainder of the present paper only the limiting case is treated where all impurity centers are either completely associated or completely dissociated. In Sec. II(a) the general procedure for treating the time-dependent equations is sketched, and subsequently the special case of simple ac is treated in detail. An expression for the current is obtained without the omission of any terms [Eq. (21)]; however, its evaluation requires the introduction of a simplified "zero solution." Then the complete expression for the frequency dependence of the equivalent susceptance and conductance is established (Sec. II(b)). Finally, in Sec. III, the theory is compared with the results of ac measurements performed in this laboratory on locally manufactured selenium disks. The theory explains the frequency dependence of capacitance and conductance over the range from about 2000 to 200,000 cps, if the presence of two species of ions of different mobilities is assumed. From this agreement ionic data, i.e. the concentrations of the carriers, their mobilities, diffusivities, and the thickness of the depletion layer, are derived and discussed.

THE theory which has been applied with the most success to the understanding of the rectifying action of semiconductors is the so-called boundary layer theory. It was developed in a simpler form by Mott,[1] and in a more elaborate form by Schottky and his collaborators.[2] The underlying concept is that the unipolar resistance of the contact between metal and semiconductor is mainly due to a thin layer, the depletion layer, which is practically depleted of the mobile particles, "electrons" or "positive holes," as the case may be. The state inside the semiconductor is regulated by space charge and diffusion and is affected in a significant way by the relevant boundary conditions.

This model gives the essential features of the observed characteristics though it proves not to be flexible enough to account for the finer details. Thus it was recognized by Mott and Schottky that field effects, due to the intensity of the electric field near the boundary, and to the discrete nature of the perturbing charges, had to be introduced.

An entire group of phenomena, which are not explained by the theory as it stands, is related to the behavior of rectifiers with time, in particular the comparatively slow changes in conductivity which are observed. It has been suggested by several authors that these changes may be due to an ionic component of the current, as opposed to the electronic component so far studied.

Now, a theory of conduction in polarizable media based on exactly the same principles as the Mott-Schottky theory was developed by the author as early as 1932.[3] At that time the theory was applied to explain the well-known time effects connected with polarization, and the boundary conditions were chosen in such a way as to realize this aim. It was not the object of the investigation, as carried out then, to give account of rectification, and this effect was automatically excluded by the symmetrical nature of the boundary conditions. However, it is only necessary to introduce boundary conditions of a dissymetrical kind, such as used by Mott and Schottky, to make the theory include unipolar conduction, i.e., rectification.

It was pointed out in the quoted paper that the changes in conductivity observed in polarizable media can be explained if two kinds of carriers were assumed present. They were then called ions of the first and of the second kind, the former being characterized by the fact that they could not leave the dielectric and therefore could not carry a current across the boundary, whereas those of the second kind were subject to the usual boundary conditions and could carry a current across the boundary.

It is evident that this approach to the problem can be extended to semiconductors and that these lend themselves particularly well to description by such a model. If the normal carriers (electrons or positive holes) are identified with the "ions of the second kind," and the ions in the semiconductor with the "ions of the first kind," the semiconductor will show the phenomena of polarizability and relatively slow changes with time, as desired.

In the present investigation the older theory of the author will be generalized in such a way as to include

* Part of the work reported here was performed under contract with the Department of the Army, Signal Corps.

[1] N. F. Mott, Proc. Roy. Soc. (London) **A171**, 27, 281 (1939).

[2] (a) W. Schottky, Z. Physik **113**, 367 (1939); (b) W. Schottky and E. Spenke, Wiss Veröffentl Siemens-Werken **18**, 225 (1939); (c) W. Schottky, Z. Physik **118**, 539 (1942).

[3] G. Jaffé, Ann Physik **16**, 217, 249 (1933).

rectification. From what has been said above, this amounts to the same as saying: the Mott-Schottky theory will be generalized in such a way as to include mobile ions. However, our treatment, the older one as well as the present one, is more general than the Mott-Schottky approach in one other regard. Mott assumes the chemical equilibrium in the semiconductor to be established and known, and Schottky treats of it in a very elaborate way[4] under the specific assumption that it is "quasi-thermic," i.e., that it is not modified by the passage of the current. This assumption is evidently correct if the time which the dissociation equilibrium requires to establish itself is short compared to the time during which the concentration of the various particles may change as a result of their motion (in the diffusion field and in the electric field). Thus, the assumption of quasithermic equilibrium breaks down at least when we are dealing with ac and when its period becomes comparable with the chemical time of relaxation.[5] Now it is our object to include in our fundamental equations the possibility that the applied voltage varies with time in an arbitrary way. Then it will not be correct to calculate the dissociation equilibrium independently of the passage of the current. The way of avoiding this is to start from the equations for the rate of change of each species of particles, instead of starting from the equilibrium equations (as do Schottky and Spenke). This leads, in the case of the stationary state, to equations which establish the modification of the dissociation equilibrium by the diffusion field and the electric field.

The fundamental equations will be derived in this more general form because they are meant to serve as a basis for various problems in the theory of conductance of crystals, semiconductors, and solutions. However, in the present paper two limiting cases will be treated only, and further simplifying assumptions will be introduced. It will be assumed that there are present in the semiconductor either positive holes and negative ions only (case A: defect semiconductor), or electrons and positive ions only (case B: excess semiconductor). Furthermore, it will be assumed that the impurity centers are, in case A, completely associated with electrons and, in case B, completely dissociated from electrons. Under these restrictions the problem can be treated in a general way by a method of successive approximations for a wide class of time-dependent applied voltages. However, for the sake of brevity, only the case of simple ac will be treated explicitly here.

I. THE FUNDAMENTAL EQUATIONS

We are treating the linear case only and assume the two electrodes to be in the planes $x=0$ and $x=l$. Dealing with case A first, we designate by p the concentration (i.e., number per cc) of the positive holes, by n_A and n the concentrations of the "acceptors" in the neutral and charged states, respectively. Since we assume all particles, even the neutral acceptor molecules, to be mobile, the concentrations p, n_A, and n, as well as the field intensity E, are, in general, functions of x and t.

Each time a neutral acceptor molecule associates with ("catches") an electron, an A^- ion and one positive hole are formed. This "formation process" may be considered as being of the first order, since the electron concentration in the highest band is practically constant. Hence, the number of positive holes (and A^--ions) formed per sec and cc is given by $k_1 n_A$ where k_1 is a rate constant.

Each time an A^--ion gives up its electron to a positive hole, one A^--ion and one positive hole disappear (by "recombination" of the positive hole with the electron). This recombination process is of the second order, the number of positive holes (and A^--ions) disappearing per cc per sec being $k_2 pn$ where p_2 is a second rate constant.

Besides these "chemical" changes, the changes due to diffusion and to migration in the field must be taken into account. Let b be the mobility of the positive holes and D their coefficient of diffusion, then the complete equation of balance will be

$$\partial p/\partial t = k_1 n_A - k_2 np + D\partial^2 p/\partial x^2 - b\partial(pE)/\partial x. \quad \text{(I)}$$

If b' and D' have the corresponding significance for the negative ions, we obtain in a similar way

$$\partial n/\partial t = k_1 n_A - k_2 np + D'\partial^2 n/\partial x^2 + b'\partial(nE)/\partial x. \quad \text{(II)}$$

The neutral acceptor molecules do not move in the electric field but may diffuse, hence their equation of balance is

$$\partial n_A/\partial t = -k_1 n_A + k_2 np + D_A \partial^2 n_A/\partial x^2, \quad \text{(III)}$$

if D_A is the relevant coefficient of diffusion.

Finally, the electric field, E, is subject to the Poisson equation

$$\partial E/\partial x = 4\pi\epsilon/K(p-n) \quad \text{(IV)}$$

if ϵ is the elementary charge, and K the dielectric constant of the semiconductor.

Equations (I) to (IV) are the fundamental equations of the problem as treated here. They suffice to determine the four unknowns, namely p, n, n_A, and E, provided adequate boundary and initial conditions are given (see below).

Our theory is descriptive in the sense that the relevant "constants" b, D, etc., and the rate constants k_1 and k_2 are considered as known functions of the temperature. Of course, what is known of these constants from other arguments[6] can easily be incorporated into the present theory.

It will not be necessary to write down explicitly the fundamental equations for excess semiconductors (case

[4] See reference 2(b), Sec. A.

[5] It will be seen in what follows that the assumption becomes illegitimate even in the case of dc provided the inhomogeneities of the field and of the particle densities are sufficiently large.

[6] Compare N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, New York, 1940).

B). Let now n, p, and p_D stand for the concentrations of the free electrons, the positive ions, and the neutral "donator" molecules, respectively. The relevant formation process will now be the release of an electron by a neutral donator molecule whereby one free electron and one positive donator-ion is formed (rate: $k_3 p_D$). The reverse recombination process consists in the formation of a neutral donator molecule by the capture of an electron by a positive ion (rate: $k_4 pn$).

It is easy to show that, with the definitions given above, the fundamental equations for case B are obtained from (I) to (IV) by replacing in these equations n_A, k_1, k_2, D_A by p_D, k_3, k_4, D_D, respectively.

Our fundamental equations show in what way the quasithermic equilibrium is affected by the migration of the particles in the electric field and in the diffusion field. In the stationary case the left-hand sides of Eqs. (I) to (III) become zero. However, only if either the neutral molecules cannot diffuse ($D_A=0$ in case A), or if the ions have no mobility ($D'=b'=0$ in case A), do Eqs. (II) and (III) reduce to the mass action equation $np/n_A=k_1/k_2$.

In the general case this equation will be correct only if the differential coefficients with regard to time, as well as the local gradients representing the influence of diffusion and migration in the field, are sufficiently small.

It is easy to write down the fundamental equations in more general cases, e.g., if electrons, positive holes, acceptors, and donators are present at the same time. It can be deduced then that for each species of particles which are not mobile the corresponding mass action equation is obtained for the stationary case.

As for the boundary conditions, we have to distinguish between positive holes (electrons) on the one hand and ions on the other hand. For the electronic carriers we assume the asymmetric boundary conditions as introduced by Mott and Schottky. Thus we prescribe in case A

$$p=p_{\mathrm{I}} \text{ for } x=0; \quad p=p_{\mathrm{II}} \text{ for } x=l, \tag{1}$$

and in case B

$$n=n_{\mathrm{I}} \text{ for } x=0; \quad n=n_{\mathrm{II}} \text{ for } x=l. \tag{2}$$

For the ions we retain the boundary conditions of our older theory. These prescribe that no ionic current can cross the two boundaries, i.e., in case A

$$b'nE+D'\partial n/\partial x=0, \quad \text{for } x=0 \text{ and } x=l, \tag{3}$$

and in case B

$$bpE-D\partial p/\partial x=0, \quad \text{for } x=0 \text{ and } x=l. \tag{4}$$

It is reasonable to assume that the neutral impurity centers, though mobile by diffusion, cannot cross the boundaries. Hence, we have the boundary conditions

$$\partial n_A/\partial x=0 \text{ or } \partial p_D/\partial x=0, \quad \text{for } x=0 \text{ and } x=l. \tag{5}$$

Finally, the field E has to be determined subject to the condition that the voltage between the electrodes

$$V(t)=\int^{l} E dx \tag{6}$$

is a prescribed function of time, related to the applied voltage $U(t)$.[7]

The initial conditions, together with Eq. (6), determine the nature of the problem in question.

A few words should be said about the physical significance of the boundary conditions as stated. As far as the electronic component is concerned, there will be a "depletion layer" at the electrode, say at $x=0$, if the values of p_{I} or n_{I} (respectively, in cases A and B) are smaller than the undisturbed value which would hold at large distances from the electrode. However, the character of the depletion layer will differ both from that of Mott (no charges whatever), and that of Schottky (no electronic charges and constant ionic density). Since it is assumed that the ions are mobile but cannot carry their charges across the boundaries, the space charge distribution will vary under the influence of the applied voltage. In the dc case, and for the high resistance direction, the ions will run against the electrode until the ionic current is stopped. This will cause a space charge opposed in sign to the electronic component and, therefore, the flow of the latter will be enhanced. In the low resistance direction the electronic component will be decreased and the "infinite catastrophe" of the Schottky theory will be avoided.

The same situation will hold with ac as long as the frequency is so low that the average time which the ions take to travel over the distance l by diffusion is small compared to the period. If the frequency becomes so high that the reverse is true, the ionic flow is no longer impeded by the boundary layer, and the ions can contribute to the conductivity according to their density and mobility. This contribution can be appreciable because in the very nature of the depletion layer the density of the ions is high while that of the electronic carriers is low. It will be seen (see Sec. III) that in selenium this effect is so large that ionic effects completely dominate electronic effects at frequencies above a few thousand cps.

Finally, it should be pointed out that the ionic boundary conditions in the form (3) and (4) represent an ideal limiting case. Actually the discharge at the electrodes will be a time-dependent process. The theory here presented has been generalized in the direction that the discharge across the boundaries is not altogether prohibited, but is regulated by a rate constant and the concentration at the electrode. Then there will be the possibility of an ionic transport across the

[7] Since we are not going to treat the dc case here, the explicit relation between $V(t)$ and $U(t)$ will not be required. It is given in case (A) by $V=U-V_s$ and in case (B) by $V=U+V_s$, where V_s represents Schottky's "diffusion potential" (see reference 6, pp. 70 and 168, and references 2(b) and 2(c)).

boundary, and its amount will depend on the joint actions of the diffusion and the discharge process. Thereby the theory becomes applicable to the case of electrolytic solutions.[8] However, this refinement of the theory is not required for semiconductors, at least not for the case of selenium here studied. The experimental results indicate that the limiting case represented by Eqs. (3) and (4) represents a sufficient approximation.

II. COMPLETE ASSOCIATION (DISSOCIATION)

(a) General Solution

Under this assumption the fundamental equations simplify considerably since all impurity centers are supposed to be ionized at all times. In order to make the formulations less cumbersome we shall treat only case A explicitly, and shall state the results for case B later.

In case A we have $n_A=0$ by assumption, hence it follows from Eq. (III) that $k_2=0$, i.e., the rate constant for recombination, must be negligibly small.

Under these circumstances the fundamental equations reduce to

$$\partial p/\partial t = D\partial^2 p/\partial x^2 - b\partial(pE)/\partial x, \tag{Ia}$$

$$\partial n/\partial t = D'\partial^2 n/\partial x^2 + b'\partial(nE)/\partial x, \tag{IIa}$$

$$\partial E/\partial x = (4\pi\epsilon/K)(p-n). \tag{IVa}$$

These equations are not linear and, therefore, can be treated only by successive approximations. If it is assumed that the voltage $V(t)$ is a periodic function which can be represented in the form[9]

$$V(t) = \sum_{m=0}^{\infty} V_m \exp(im\omega t), \tag{7}$$

Eqs. (Ia), (IIa), and (IVa) can be treated in a general way. If expressions similar to (7) are assumed for p, n, E, and the current density j, the differential equations are linearized. They can be solved in a formal way if the field is supposed to be known, and this leads to an alternate procedure for the actual determination of the unknowns similar to the one given below for a more special case.

We have carried through the calculations using infinite series as indicated,[10] however, for the sake of brevity we will restrict ourselves here to the simpler case where the series (7) reduces to its first two terms.

Even then the complete solutions for p, n, E, and j involve infinite series of the type (7), owing to the nonlinearity of the differential equations. This means physically that a semiconductor described by the present model will act as a nonlinear element, i.e., will respond to a simple harmonic emf not only with the same period but also with all its overtones. Again, since we are interested mainly in the simple ac response, we shall disregard all terms involving m higher than one.

Consequently, we assume not only that Eq. (7) reduces to two terms, but also that all unknowns are of the form

$$p = p_0 + p_1 \exp(i\omega t). \tag{8}$$

The case $V_1=0$ has to be treated separately. It represents the dc characteristic which has been developed to the second approximation.[11] However, since no new observations on the dc characteristic are available, we shall not treat the dc case explicitly here.

If the expressions for p, n, and E of the form (8) are substituted into Eqs. (Ia) and (IIa), these reduce to

$$i\omega p_1 - D(d^2p_1/dx^2) = -b(dF_1/dx) \tag{9}$$

and

$$i\omega n_1 + D'(d^2n_1/dx^2) = b'(dF_1'/dx), \tag{10}$$

where we have set

$$F_1 = E_0 p_1 + E_1 p_0, \tag{11}$$

$$F_1' = E_0 n_1 + E_1 n_0. \tag{12}$$

If we treat F_1 and F_1' as known, we obtain the following solutions for p_1

$$p_1 = \mathcal{E}_1 \exp(\mu, x) + \mathfrak{F}_1 \exp(-\mu_1 x) + (b/2D)[G_1(x) + \bar{G}_1(x)] \tag{13}$$

and n_1

$$n_1 = \mathcal{E}_1' \exp(\mu', x) + \mathfrak{F}_1'(-\mu^1, x) - (b'/2D') \times [G_1'(x) + \bar{G}_1'(x)]. \tag{14}$$

Here we have set

$$\mu_1 = \rho_1(1+i), \quad \rho_1 = (\omega/2D)^{\frac{1}{2}}, \tag{15a}$$

$$\mu_1' = \rho'(1+i), \quad \rho_1' = (\omega/2D')^{\frac{1}{2}}, \tag{15b}$$

and

$$G_1(x) = \exp(\mu_1 x) \int_0^x \exp(-\mu_1 y) F_1 dy, \tag{16a}$$

$$\bar{G}_1(x) = \exp(-\mu_1 x) \int_0^x \exp(\mu_1 y) F_1 dy. \tag{16b}$$

The definitions for G_1' and $\bar{G}_1'$ are obtained from (16a), (b) by substituting μ_1', $-F_1$, for μ_1, F_1. The constants of integration $\mathcal{E}$ and $\mathfrak{F}$ remain to be determined.

[8] This part of the theory and its application to electrolytic solutions will be published later.

[9] Since, for the sake of simplicity, the summation in Eq. (7) is not extended to $-\infty$, Eq. (7) does not represent the most general case. It holds if $V(t)$ is an arbitrary function of time which is either even or odd.

[10] Final report of the Signal Corps Project No. 152 B.

[11] See reference 10. In that report the theory is applied to the slow changes in time which are observed upon prolonged application of the direct voltage. The results are in qualitative agreement with observations, including the occasionally observed appearance of a current maximum, some time after the application of the field. However, the ionic mobilities necessary to explain such an event come out to be unreasonably large (of the order of $\frac{1}{10}$ of the electronic mobility), if a quantitative fit is attempted.

If now Eqs. (13) and (14) are substituted into the Poisson equation, an integral equation for E_1 is obtained which we shall indicate only in the abbreviated form

$$E_1 = V_1/l + E_1^+ + E_1^- . \tag{17}$$

The expressions for E_1^+ and E_1^- are somewhat lengthy[10] and need not be given here since they are not going to be used. They represent the space charge fields arising from the holes and negative ions, respectively, while V_1 is the applied ac voltage.

We are mainly interested in the current which will flow for a given value of V_1. The total current j_{tot} consists of the displacement current

$$j_{dis} = (K/4\pi)\partial E/\partial t$$

and the convection current

$$j_{conv} = e[(bpE - D\partial p/\partial x) + (b'nE + D'\partial n/\partial x)]. \tag{18}$$

The resultant value depends on the constants of integration which, in turn, depend on the boundary conditions. In case A the density p refers to positive holes and must be subjected to the conditions in Eq. (1). Since p_I and p_{II} are supposed to be independent of time, the boundary conditions must be realized in the dc component; consequently, we must demand $p_1 = 0$ for $x = 0$ and $x = l$. This determines $\mathcal{E}_1$ and $\mathcal{F}_1$. On the other hand, n_1 refers to negative ions. Hence the boundary conditions (3) will have to be used, which determines $\mathcal{E}_1'$ and $\mathcal{F}_1'$ to be

$$\mathcal{E}_1' = \mathcal{F}_1' = (b'/2D')[G_1'(l) - \bar{G}_1'(l)]/ [\exp(\mu_1' l) - \exp(-\mu_1' l)]. \tag{19}$$

The total current can be obtained in a variety of ways. The knowledge of E_1^+ and E_1^- is not required if the following procedure is adopted. Since the total current is the sum of the displacement current and of the convection current, and since $\partial E/\partial t$ reduces under our assumptions to ∂ ∂t, we obtain by simple integration

$$(j_1)_{tot} = i\omega(K/4\pi)V_1/l + (1/l)\int_0^l j_{conv}\,dx, \tag{20}$$

where j_{conv} is given by Eq. (18). The integral is easily transformed if it is taken into account that p_I vanishes at $x = 0$ and $x = l$, and if the expressions (11), (12), (14), and (19) are used. The result is given by

$$(j_1)_{tot} = i\omega(K/4\pi)(V_1/l) + (eb/l)\int_0^l F_1\,dx + (eb'/l)\int_0^l F_1'\,dx - (eb'/l)[G_1'(l) + \exp(\mu_1' l)\bar{G}_1'(l)]/ [\exp(\mu_1' l) + 1]. \tag{21}$$

In case B an expression can be derived which is identical with (21) except that in the last term b, μ_1, G_1, $\bar{G}_1$ replace b', μ_1', G_1', $\bar{G}_1'$.

Our equations (13), (14), (17), and (21) for the densities, the field and the current density are rigorous consequences of the differential equations (9), (10), and the corresponding Poisson equation. However, they are formal solutions only, in the sense that they represent rather complicated integral equations for the unknowns. The value of this formal solution consists in the fact that it leads the way to an actual calculation of the unknowns by successive approximations. The procedure is as follows: we begin by assuming any reasonable zero solution for p_1, n_1, and E_1. In the simplest case, when $V_0 = 0$ and V_1 is small, a homogeneous field and no free space charge would represent such a zero solution. By the aid of this zero solution improved first-order solutions for p_1 and n_1 may be obtained from (13) and (14), and then for the field from (17), or from the Poisson equation.

In this way the procedure goes on by the alternate use of Eqs. (13) and (14) on the one hand, and of Eq. (17) on the other hand. Finally, the current density is obtained from Eq. (21). The value of the procedure will, of course, depend greatly on the choice for the zero solution. In the case of dc, only the field E_0 has to be assumed. The assumption of a constant field leads to Mott's formula, the assumption of a linear field to an improved Schottky formula which avoids the "infinite catastrophe." The zero approximation will be "adequate" if the next step in the alternate procedure does not modify it significantly.

b. The ac Characteristic

The second term of (21) represents the electronic contribution. It is not our object here to study the frequency dependence of the impedance as far as it is due to the electrons or positive holes, though our method is applicable. The ac behavior of the electronic component of semiconductors has been studied by Spenke,[12] who obtained results which are explicit for small frequencies. Furthermore, Schottky[13] has shown that the "internal time of relaxation" for semiconductors, as far as the electronic component is concerned, is very short, of the order of 10^{-10} sec. Therefore, up to frequencies almost as high as 10^{10} cps, the behavior of the electronic component can be described by equivalent values of the capacitance and resistance which are constants. On the other hand, Eq. (21) indicates a frequency dependence for much lower frequencies due to the ionic component.

For these reasons we shall assume that the frequency is limited to such values that the first two terms of Eq. (21) can be written in the form

$$(j_1)_{el} = (1/R_{el} + i\omega C_{el})V_1, \tag{22}$$

where C_{el} and $1/R_{el}$ are the constant effective values of the electronic contribution to the parallel capacitance and conductance, respectively.

[12] W. Spenke, Wiss. Veröffentl Siemens-Werke **20**, 40 (1941).
[13] W. Schottky, reference 2, p. 550 ff.

The last two terms of (21) represent the ionic contribution to the current. It will be seen that an exact evaluation would require the knowledge of E_0 and n_0, i.e., the solution for the dc case. If we should attempt the evaluation using the solutions obtained for this case,[10] the calculations would become prohibitively complicated. Therefore, we are going to introduce simpler approximations with the double object of simplifying the integrations and of making the knowledge of the dc solution unnecessary. This object is achieved by the simple assumption that in calculating the ac current, the inhomogeneity of the ionic distribution may be disregarded, though the inhomogeneity of the field should be retained.

Consequently we assume, as zero approximation for n,

$$n_0=c_0, \quad n_1=0, \tag{23}$$

where c_0 is the undisturbed value.

A few words should be said regarding the validity of these assumptions. If we suppose that the applied ac voltage V_1 is sufficiently small, our second assumption (23) becomes legitimate since a weak V_1 will not contribute noticeably to the space charge distribution. Nor is it inconsistent to retain the inhomogeneity of the field as, in the case of small voltages, it is much more marked than the inhomogeneity of the ionic distribution.[14]

The first assumption (23) is less justified and would be quite inadequate if the case of an applied bias were to be studied. Therefore, we have to assume that there is no applied dc voltage present. Even then there may be a space charge distribution of the negative ions near the electrodes. Its value will depend strongly on the boundary conditions which we impose on n, and will affect the current indirectly. However, it is easy to see that this influence is limited to low frequencies. The discussion given in Sec. I shows that the conditions near the electrodes become of less and less importance as the frequency increases. Therefore, the first assumption (23) becomes adequate as those values of the frequency are reached where the ionic current begins to rise strongly from its low value at low frequencies. In the applications we are going to study this region exclusively.

As for the field strength, our second assumption (23) makes the knowledge of E_0 unnecessary. For E_1 we assume

$$E_1=2C_1\cosh[\alpha(x-l_1)], \quad l_1=l/2, \tag{24}$$

with α given by

$$\alpha^2=(8\pi\epsilon/K)(b'/D')c_0. \tag{25}$$

This form of the field is suggested by the fact that it represents a first-order approximation in the static case studied in the author's older theory.[15] Furthermore, it can be proven that it is an "adequate solution" in the dc case, even when a current flows.[10] This, of course, does not prove yet that Eq. (24) is a good approximation in the ac case, except for low frequencies. As a matter of fact, it can be foreseen that E_1 becomes more and more homogeneous as the frequency increases.[16] From this point of view it might appear that the simpler assumption $E_1=$ const represents as good a zero approximation as (24). However, if $E_1=$ const is used, the possibility of determining the thickness of the polarization layer (see below) is lost. For this reason we retain the form (24).

Finally, we assume that there is only one depletion layer, at $x=0$, but no polarization layer at the other electrode. This we realize by using the solution (24) only from $x=0$ to $x=l/2=l_1$. Then the constant C_1 becomes

$$2C_1=\alpha V_1/\sinh(\alpha l_1) \tag{26}$$

if V_1 is the voltage applied to the layer of thickness l_1.

In calculating the current, we may still use Eq. (21). Owing to the complete symmetry of our assumptions with regard to $x=l/2$, Eq. (21), when divided by $2V_1$, yields the impedance of two polarization layers in series, and therefore, half of the conductance and capacitance of each of them.

If we substitute (23) to (26) into Eq. (21) the integrations become quite simple and yield

$$(j_1)_{\text{ion}}=\epsilon b'c_0(V_1/l_1)\{1-(1/(\alpha^2-\mu_1'^2))\times[\alpha^2-\alpha\mu_1'\nu\tanh(\mu_1'l_1)]\}. \tag{27}$$

Here the abbreviation

$$\nu=\coth(\alpha l_1) \tag{28}$$

has been introduced. Since μ_1' [see Eq. (15b)] is complex, Eq. (27) defines a complex admittance. If the real and imaginary parts are separated, the admittance for the layer extending from $x=0$ *to* $x=l_1$ assumes the form

$$Y_1=1/R_{s1}+(1/R_\infty)\Gamma_2(\lambda',\beta)+i\omega[C_{s1}+C_0\Gamma_1(\lambda',\beta)]. \tag{29}$$

The two functions of two arguments, Γ_1 and Γ_2, are defined by

$$\Gamma_1(\lambda',\beta)=(1+4(\lambda'/\beta)^4)^{-1}[(\nu/\lambda')g_1(\lambda')-2/\beta+(2\nu\lambda'/\beta^2)g_2(\lambda')] \tag{30}$$

and

$$\Gamma_2(\lambda',\beta)=(1+4(\lambda'/\beta)^4)^{-1}[(\nu\lambda'/\beta)g_2(\lambda')-2\nu(\lambda'/\beta)^3g_1(\lambda')+4(\lambda'/\beta)^4], \tag{31}$$

and the two functions of one argument, g_1 and g_2, are

$$g_1(\lambda')=(\sinh\lambda'+\sin\lambda')/(\cosh\lambda'+\cos\lambda') \tag{32}$$

and

$$g_2(\lambda')=(\sinh\lambda'-\sin\lambda')/(\cosh\lambda'+\cos\lambda'). \tag{33}$$

[14] Using the static solution and the constants obtained in Sec. III, we find for the ratio of n_1 at the anode to the undisturbed value about 1.5 for a potential difference of 10^{-2} volt, whereas the same ratio for E is of the order of 100.

[15] See reference 3, pp. 228 to 229.

[16] H. Chang has carried the calculation of E_1 to the second order and verified the statements in the text.

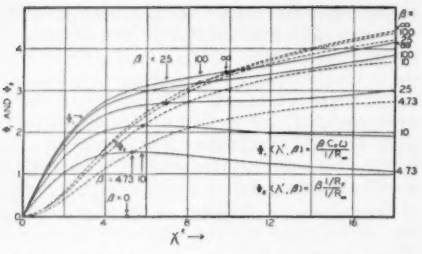


FIG. 1. Theoretical curves representing the reduced susceptance, $\Phi_1(\lambda', \beta)$, (full lines), and the reduced conductance, $\Phi_2(\lambda', \beta)$ (broken lines), as functions of the reduced frequency, λ'^2, for various values of the reduced plate distance β.

The variable λ' is given by

$$\lambda' = 2l_1(\omega/2D')^{\frac{1}{2}}, \tag{34}$$

and the dimensionless parameter β by

$$\beta = 2\alpha l_1. \tag{35}$$

The variable λ' is dimensionless, but the combination $(2l_1)^2/D' = \tau'$ has the dimension of time and may be taken as the "time of relaxation" for the ions of diffusivity D'. It represents the average time such ions take to travel over the distance $2l_1$ by diffusion. Consequently, $(\lambda')^2$ itself is proportional to the ratio of τ' to the length of the period. It is seen that the time-dependence of Y_1 is expressed by this ratio exclusively. This specifies quantitatively the statements made in the introduction.

As for the dimensionless parameter β, it was introduced in the author's older theory by the name of "reduced plate distance" and designated by δ (see reference 3, pp. 226 ff.). It characterizes the behavior of the polarized layer. If β is sufficiently large, and that means from about $\beta = 10$ on, the thickness of the polarization layer will be small compared with the total thickness of the dielectric.

Since Γ_1 tends towards unity as $\lambda' \to 0$ and $\beta \to \infty$, C_0 is the equivalent parallel capacitance (of ionic origin) in that limiting case. Its value is

$$C_0 = \epsilon b' c_0/(\alpha D'). \tag{36}$$

On the other hand, Γ_2 tends to unity as $\lambda' \to \infty$. Hence

$$1/R_\infty = \epsilon b' c_0/l_1, \tag{37}$$

defines the equivalent parallel conductance (of ionic origin) in the limiting case of very high frequencies. All constants refer to one cm³ area.

If we designate by R_p and C_p the equivalent parallel resistance and capacitance, respectively, their frequency dependence will be expressed by

$$C_p = C_0\Gamma_1(\lambda', \beta), \quad 1/R_p = (1/R_\infty)\Gamma_2(\lambda', \beta). \tag{38}$$

In order to illustrate this dependence, we have plotted in Fig. 1 the two dimensionless functions

$$\begin{aligned} \Phi_1(\lambda', \beta) &= \beta\omega C_p/(1/R_\infty), \\ \Phi_2(\lambda', \beta) &= \beta(1/R_p)/(1/R_\infty). \end{aligned} \tag{39}$$

They represent reduced values of the susceptance and conductance. As abscissas we have chosen λ'^2 which is proportional to the frequency, and representative values of β, from $\beta = 4.73$ to $\beta = \infty$, have been selected.

The curves for Φ_1 all pass through a maximum and ultimately approach zero with $\omega \to \infty$. These features are not exhibited by the diagram as they occur for higher values of λ'^2. On the other hand, the curves representing Φ_2 (broken lines) increase to the asymptotic value β. Thus, with increasing frequency the ionic conductance increases from zero to the maximum value given by (37). Conversely the capacitance decreases from its initial value at $\omega = 0$, $C_0(1 - 2/\beta)$ to zero.

For a given value of β the corresponding curves for Φ_1 and Φ_2 intersect at a characteristic frequency $f = f^*$, and this point of intersection moves from $\lambda'^2 = 5.1$ to $\lambda'^2 = 9.8$ as β varies from 0 to ∞. It will be seen in the next section that such points of intersection have been observed and play an important part in the evaluation of the ionic constants.

The formulas so far developed in this section refer to defect semiconductors. For excess semiconductors analogous results are obtained by replacing in (21), and correspondingly in all subsequent equations, the symbols p, n, b', D', μ_1' by n, p, b, D, μ_1, respectively.

The thickness of the layer l_1 over which the solution is to be applied, remains to be determined. This we do by taking into account the bulk resistance of the entire semiconductor, and by requiring the current and the field to be continuous at $x = l_1$. By this procedure it can be shown[10] that

$$\beta = 2\alpha l_1 = 4\ln\alpha - 2\ln(\omega/2D'), \tag{40}$$

if λ' is small and β not small (say $\beta > 10$).

Thus l_1 is, for a given concentration of the ions, a slowly variable function of the frequency. As λ' increases, the approximate solution (40) ceases to be correct, and graphical methods have to be applied (see Sec. III); however, l_1 continues to decrease slowly with increasing λ'. On the other hand, with decreasing λ' the thickness l_1 will increase until it becomes equal to the total thickness of the semiconductor.

In the applications it will become imperative to assume the presence of more than one species of moveable ions. The resulting expression for the admittance is of the same form as (29); only there are as many terms involving functions $\Gamma_{1i}(\lambda_i', \beta)$ and $\Gamma_{2i}(\lambda_i', \beta)$ as there are different species of mobile carriers (positive or negative). Naturally, all of them move in the same field, and this necessity finds its expression in the fact that β is the same throughout. Provided all ions have the same valence, the expression for α has to be changed

from (25) to

$$\alpha^2=(8\pi\epsilon/K)(b/D)\sum_i c_i, \qquad (41)$$

where c_i is the concentration of the i-species of carriers and where the summation extends over all carriers of one sign.[17]

III. COMPARISON OF THE THEORY WITH EXPERIMENTAL RESULTS

It is well known that the ac characteristic of semiconductors shows frequency dependence at much lower frequencies than might be expected from the "internal time of relaxation" of electronic carriers. This was one of the main reasons for developing the present theory. At the same time an extensive series of bridge measurements was carried through.[18] These measurements cover a range of frequencies from 10 cps to 200,000 cps and were taken on locally fabricated selenium disks between nickel and cadmium electrodes.

It has been customary, particularly in the work of Schottky and his associates,[19] to discuss the ac behavior of semiconductors in terms of equivalent circuits. Valuable as this method is for a rapid survey of results, it cannot be considered adequate for the description of the finer details. If the fundamental concepts of the present investigation are accepted, Eq. (29) and its generalization for several species of carriers indicate clearly enough that the equivalent parallel susceptance and conductance are rather complicated functions of frequency. Furthermore, the more general treatment indicated at the beginning of Sec. II(b) proves that a simple harmonic impressed emf will cause overtones of all orders in the current. Therefore, in our opinion, the correct description of a semiconductor is that by numbers and mobilities of the relevant carriers. An attempt at determining these from the measurements mentioned above is given in what follows.

The experimental curves representing ωC_p and $1/R_p$ as functions of frequency invariably show the character of the theoretical curves given in Fig. 1, though the numerical values may vary considerably from one disk to another and for one and the same disk with the previous treatment. It was pointed out in Sec. II(b) that there is a point of intersection between the ωC_p and the $1/R_p$ curves at a determined frequency f^*. The determination of ionic constants starts most conveniently from these points where they have been observed.

In Table I we have given all cases where such points of intersection have been observed between about 100 cps and 200,000 cps. In some cases a point of intersection was observed before and after formation, where the formation consisted in the application of 40 to 160 volts in the blocking direction over a period of 3 to 6 hours. Many other disks indicated the existence of a point of intersection at higher frequencies than were observed.

The effective resistances of the locally fabricated disks at the lowest observed frequencies, i.e., 10 cps, were considerably (from 30 to 100 times) larger than the dc resistance of commercial disks. Thus, it is evident that the locally fabricated disks must have contained much less accidental impurities than the commercial ones.

Furthermore, it is characteristic of the locally fabricated disks that the electronic component may be disregarded altogether for the interval of frequencies over which a match with the theory was attempted. This is justified by the fact that, at lowest frequencies, where only the electronic component can contribute to the observed conductances, the conductance is negligibly small compared to that in the matched interval (see Figs. 2 and 3). Therefore, in the following computations the electronic component of the current has been disregarded.

It will be noticed from Table I that disk 74, in its two states, represents extreme conditions. (The forming in this case was performed by application of 40 volts for 4 hours.) In all cases the frequencies f^* and the corresponding values of ωC_p and $1/R_p$ are contained within the limits set by the disk 74 (with one unimportant exception in the case of disk 61). For this reason the laborious calculations were carried through only for the two states of disk 74, which both refer to zero bias (i.e., $U_0=0$). If the very considerable changes occurring in disk 74 during formation are represented in a satisfactory way, it seems reasonable to assume that also the other, intermediate, cases can be matched.

A great simplification of the numerical evaluations arises from the fact that the values of λ' and β which belong to f^*, say λ'^* and β^* are completely determined by the theory, as long as only one species of ions is being considered. The intersection occurs for

$$\Phi_1(\lambda',\beta)=\Phi_2(\lambda',\beta). \qquad (42)$$

If this equation is combined with one which arises from the boundary conditions for $x=l_1$, there remain two equations for the determination of λ'^* and β^*. We have

TABLE I. Observed points of intersection between the curves representing susceptance (ωC_p) and conductance ($1/R_p$).

Disk	Virgin f^* (cps)	Virgin $1/R_p=C_p\omega$ (μmhos)	Formed f^* (cps)	Formed $1/R_p=C_p\omega$ (μmhos)
30	1.18×10^3	2.29×10^4	not observed	
47	0.312×10^3	2.03×10^4	not observed	
61	0.652×10^3	2.10×10^4	0.472×10^3	1.07×10^4
67	1.38×10^3	4.12×10^4	not observed	
74	2.05×10^3	4.12×10^4	0.416×10^3	1.29×10^4
84	1.02×10^3	2.98×10^4	0.647×10^3	1.38×10^4

[17] These results are obtained by first generalizing the stationary case of our older theory (see reference 3, paragraph 1) so as to include several species of carriers. This leads in a simple way to a first-order field distribution given by Eq. (24) with the modified value of α just defined. All other calculations remain the same as indicated in this section, and the final result for the current, and thereby for the admittance, is obtained by summation.

[18] The bridge measurements were performed by Mr. Allan McDonald. The details of the method and of the procedure in fabricating the disks are to be found in Report 11 and 3rd annual report (bridge measurements), and 2nd annual report (preparation of disks), Signal Corps Project No. 152 B.

[19] W. Schottky and W. Deutschmann, Physik Z. 30, 839 (1929).

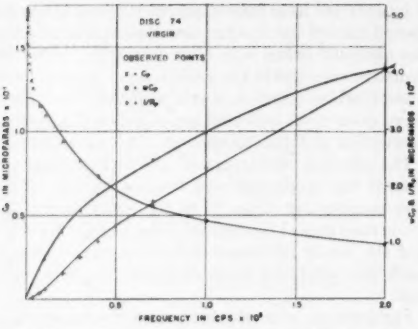


FIG. 2. Comparison of the theoretical curves for C_p, ωC_p, and $1/R_p$ with observed values. Selenium disk No. 74, virgin.

carried through the evaluation by graphical procedures; they lead to the result

$$\lambda'^{*}=2.30, \quad \beta^{*}=4.73. \tag{43}$$

With λ'^{*} and β^{*} known, C_0 and $1/R_\infty$ can be determined from the curves of Fig. 1. Then, Eqs. (25) and (34) to (37), together with Townsend's relation

$$b'/D'=\epsilon/kT, \tag{44}$$

represent a system of six simple equations for the determination of α, β, l_1, b', D', and c_0.

Unfortunately, the attempt to match experimental curves by our theoretical formulas invariably leads to the conclusion that more than one species of ions, at least two, have to be assumed as present in the semiconductor if the curves are to be matched over a wide range of frequencies. This might have been anticipated if the ions in question are of accidental origin; a whole "spectrum" of ions of different mobilities would appear probable.

In the following computations two species of ions are assumed to contribute essentially. Furthermore, β is treated as a constant, though it really is a slowly varying

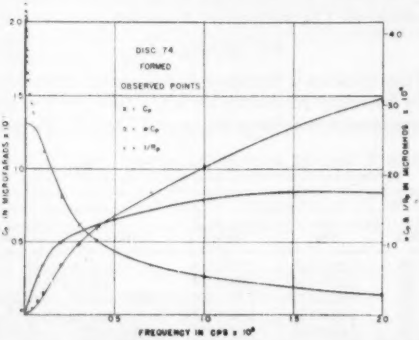


FIG. 3. Comparison of the theoretical curves for C_p, ωC_p, and $1/R_p$ with observed values. Selenium disk No. 74, formed.

function of the frequency. Naturally β may change more markedly by the formation process which affects the number of ions present, and possibly their mobilities, and this circumstance has to be taken into account.

If there are two species of ions, it follows from the last paragraph of Sec. II(b) that the effective values of ωC_p and $1/R_p$ will be given by

$$\omega C_p=A_1\Phi_1(\lambda_1', \beta)+A_2\Phi_1(\lambda_2', \beta) \tag{45}$$

and

$$1/R_p=A_1\Phi_2(\lambda_1', \beta)+A_2\Phi_2(\lambda_2', \beta), \tag{46}$$

TABLE II. Experimental constants used in the calculation of the curves for C_p, ωC_p and $1/R_p$ drawn in Figs. 2 and 3.

Disk	β	A_1 (μmhos)	A_2 (μmhos)	τ_1 (sec)	τ_2 (sec)
74, virgin	3.80	2.96×10^{4}	5.92×10^{3}	2.78×10^{-6}	2.52×10^{-5}
74, formed	4.73	0.87×10^{4}	4.96×10^{3}	4.80×10^{-6}	3.60×10^{-5}

where Φ_1 and Φ_2 are the functions defined by Eq. (39) and plotted in Fig. 1.

The experimental curves have to be matched with (45) and (46). It should be stressed that the curves for ωC_p and $1/R_p$, being mutually related, should be matched by the same constants, and if this is feasible it represents a strong argument in favor of the interpretation given here.

There are five constants available in Eqs. (45) and (46), namely A_1, A_2, β, and two constants which reduce the frequency scale to the λ'^{2}-scale for the two species of ions. Since λ'^{2} is connected with the frequency by its definition (34), the "internal time constants"

$$\tau_i=l_1^{2}/D_i', \quad i=1, 2, \tag{47}$$

may be used as such constants.

We have determined the five constants by trial and error using the following procedure. First one species of ions is considered only. By making the theoretical point of intersection coincide with the observed one, the three relevant constants (say A_1, β, and τ_1) are found without ambiguity, but the curves do not match sufficiently

TABLE III. Limiting values of specific conductivity, σ_{i_0}, and capacitance, C_{0i}, as derived from the experimental data of Table II.

Disk	σ_{10} (μmhos/cm)	σ_{20} (μmhos/cm)	C_{01} (μfd/cm²)	C_{02} (μfd/cm²)
74, virgin	3.24×10^{4}	0.649×10^{4}	1.23×10^{-2}	2.23×10^{-2}
74, formed	1.19×10^{4}	0.677×10^{4}	0.72×10^{-2}	3.07×10^{-2}

well. Then, a second (slower) species of ions is introduced and the constants first determined are modified until the matching is as good as feasible with the five constants available.

The choice of the parameter β should be restricted to narrow limits. It will be seen from Fig. 1 that the shape of the theoretical curves changes but slowly with the value of β; hence, β cannot be determined with any degree of accuracy. On the other hand, the mobilities (and diffusivities) change strongly with β (being propor-

TABLE IV. Ionic constants, and thickness of the depletion layer l_1, derived from the data of Tables II and III.

Disk	l_1 (cm)	D_1' (cm²/sec)	b_1' (cm²/volt×sec)	c_1 (cm⁻³)	D_2' (cm²/sec)	b_2' (cm²/volt×sec)	c_2 (cm⁻³)
74, virgin	0.79×10^{-5}	4.5×10^{-5}	1.71×10^{-3}	0.94×10^{15}	5.0×10^{-4}	1.90×10^{-2}	1.70×10^{16}
74, formed	1.04×10^{-5}	4.5×10^{-5}	1.71×10^{-3}	0.45×10^{15}	6.0×10^{-4}	2.27×10^{-2}	1.92×10^{16}

tional to the square of β). Hence, β should not vary very much. In order to introduce as little ambiguity as possible into the numerical evaluations, the following procedure for the determination of β was adopted. For the formed disk the theoretical value $\beta=4.73$, which is valid for a unique group of ions, was chosen because the calculations showed that for the formed disk the one group of ions (the slower one) predominates rather strongly. To represent, then, the virgin disk, β had to be decreased, but this was done as little as possible.

The degree of agreement which could be realized over the range of frequencies from about 2×10^3 to 2×10^6 cps will be evident from Figs. 2 and 3 where the curves are calculated from theory. The constants used in these computations are given in Table II, and Table III contains the limiting values (per cm² area) of capacitance (as $f\to0$), C_{0i}, and of specific conductance (as $f\to\infty$), σ_i. These limiting values have been obtained by extrapolation on the basis of the theoretical curves. The calculations involve the area of the disks utilized, which was 3.47-cm², and of the dielectric constant of selenium, which was found in the literature to be $K=6.3$.

The agreement over the indicated range of frequencies is about as good as might be expected in consideration of the various approximations which had to be introduced, as well into the theory, as into the numerical evaluations. At lower frequencies systematic deviations set in. The curves for C_p have been added because they show more clearly than those for ωC_p that for lower frequencies ($f<2\times10^3$ cps) a new group of ions, of still lower mobility, makes its appearance. It would not be difficult to represent this behavior by adding a third group to the theoretical representation.

Finally, the experimental constants of Table II, together with the extrapolated values of Table III, permit of determining all ionic constants if the reasonable assumption $p_{II}=c_0=c_1+c_2$ is made. It means physically that the number of mobile negative ions (in the undisturbed state) is equal to the number of positive holes. The following equations are available: two equations of form

$$\sigma_i=\epsilon b_i' c_i, \quad i=1, 2, \tag{48}$$

(from Eq. (37)), two equations from (44), two equations of form (47), and the one equation which results from the substitution of (41) into the definition of β, (35). These are seven equations for the unknowns c_i, b_i', D_i' ($i=1, 2$), and l_1.

The values of the ionic constants, including l_1, obtained in this manner are contained in Table IV. The concentrations of the mobile ions, c_1 and c_2, appear quite reasonable, and the thickness of the depletion layer, l_1, is of the same order of magnitude found by other observers. The mobilities and diffusivities are of particular interest since the latter may be considered as determined directly from a comparison with known frequencies. The mobilities are distinctly of electrolytic order of magnitude, the swifter ions having a mobility slightly above that of simple electrolytic ions (the mobility of the K^+-ion is $b=6.8\times10^{-4}$ cm²/(volt×sec) at 18°C), and the slower ions having somewhat smaller mobilities. In all events, these mobilities are far too small to be attributed to the positive holes. On the other hand, the values found here appear somewhat high for ions in a solid medium. This might very well be due to the lack of accuracy in the determination of β (see above).

With regard to the formation process the only conclusion which can be drawn with some certainty from the numbers of Table IV is that the importance of the swifter group of ions decreases by the application of the forming potential. The ratio c_1/c_2 decreases from 0.54 to 0.23, and this result persists even when the experimental constants A_1 and A_2 are determined in a slightly different way (within the limits which will not seriously affect the agreement between experiment and theory). The decrease of c_1 is easily interpreted; it indicates that the fastest ions are electrolyzed out first by the formation process.

On the whole, we can summarize the analysis of the experimental data by stating that the present theory coordinates the observed facts in a satisfactory way. In all events it shows the importance of the ionic component in the conductance of semiconductors like selenium, since the facts correlated here cannot be explained on the basis of a purely electronic theory.

The usefulness of the theory developed in the present paper is not limited to semiconductors. Either directly or with slight modifications, it is applicable to all cases where polarization effects are caused by the motion of ions. Work on such applications to electrolytes and to poorly conducting liquids is in progress and will be reported on separately.

ACKNOWLEDGMENTS

The author wishes to express his sincere obligation to the project engineer, Dr. L. W. Morris, and to Mr. A. McDonald for letting him utilize and publish here results obtained by them in the course of their work on the Signal Corps Project No. 152 (B). He also is indebted to Dr. L. W. Morris for many clarifying discussions, and to Dr. K. Lehovec for valuable suggestions.

Letters to the Editor

PUBLICATION of brief reports of important discoveries in physics may be secured by addressing them to this department. The closing date for this department is five weeks prior to the date of issue. No proof will be sent to the authors. The Board of Editors does not hold itself responsible for the opinions expressed by the correspondents. Communications should not exceed 600 words in length.

Transition Effect for Neutron Production in Lead*

S. B. TREIMAN† AND W. FONGER
Institute for Nuclear Studies, University of Chicago, Chicago, Illinois
(Received November 23, 1951)

THE production rate of cosmic-ray stars in photographic emulsions placed under various thicknesses of lead has been studied by a number of investigators,[1-8] and the existence of a transition maximum, occurring within the first few centimeters of lead, has been observed. A similar result has been obtained for single protons in emulsions.[9] In this note we report on a set of measurements dealing with a closely related problem: The absorption in lead of the neutron-producing radiation. This problem has been investigated at airplane altitudes by Simpson,[10] and a transition maximum at 15–18 g/cm² of lead has been observed. The measurements reported here were performed at Chicago (750 mm Hg average pressure).

The experimental arrangement is shown in Fig. 1. The aim of the experiment was to measure the production rate of neutrons in a fixed thickness (½ in.) of lead "producer," as a function of lead absorber thickness. The neutron detector consisted of a tray of three BF_3 proportional counters connected in parallel and embedded in a large block of paraffin (19×30×48 in.), which serves as a neutron moderator. We denote the counting rates by $R(x, p)$ and $R(x, 0)$, where x is the thickness of absorber and where the symbols p and 0 signify, respectively, the presence or absence of producer. The counting rates were normalized to a constant counting rate on a monitoring apparatus similar to the absorption apparatus shown in Fig. 1, but with a different (and fixed) arrangement of lead. The pressure coefficients for the two sets of apparatus were found to be the same, irrespective of absorber (or producer) thickness. This normalization procedure also corrects for possible time variations in cosmic-ray intensity.

In the absence of producer, the total counting rate $R(x, 0)$ contains a term $A(x)$ due to neutrons produced in the absorber, a term $W(x)$ due to neutrons produced in the paraffin block, and a term B due to "background" (e.g., alpha-contamination):

$$R(x, 0) = A(x) + W(x) + B. \quad (1)$$

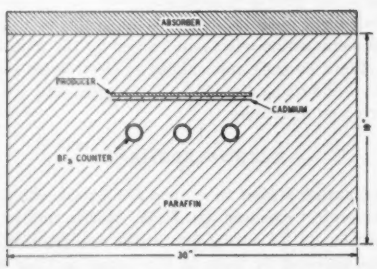


FIG. 1. Sketch of apparatus used to measure neutron production rate in half-inch lead "producer," as a function of lead absorber thickness. The cadmium sheet reduces the contribution to the counting rate from neutrons produced in the absorber.

Because of the large size of the paraffin block, the contribution from neutrons produced outside the block is negligible. The background rate was measured by covering the counters with cadmium shields, and the value $B = 3.1 \pm 0.1$ counts/min was obtained. Within this accuracy B was independent of absorber and producer thickness. When the producer is inserted, there is an additional contribution $P(x)$ due to neutrons coming from the producer. Also, a certain fraction f of the neutrons from the absorber is scattered by the producer and is no longer detected. (The contribution from the paraffin, however, is assumed to remain unchanged, because of the approximately symmetric distribution of the paraffin with respect to the BF_3 counters.) Thus,

$$R(x, p) = P(x) + (1-f)A(x) + W(x) + B. \quad (2)$$

From Eqs. (1) and (2) we obtain

$$P(x) = R(x, p) - R(x, 0) + f[R(x, 0) - W(x) - B]. \quad (3)$$

The last term in this equation is only a small correction term, particularly for small values of x. A sufficiently accurate estimate of f is obtained by assuming that the energy distribution of the neutrons produced in lead can be approximated by a Ra+Be source.[11] A 5-mC source was placed on top of the paraffin block and the reduction in counting rate, due to the insertion of the producer, was noted. The value $f = 0.060$ was obtained. The function $W(x)$, which represents the counting rate due to neutrons produced in the paraffin, is proportional to the flux of neutron-producing radiation which emerges from the absorber of thickness x. Except at small thicknesses where a transition effect occurs, we can anticipate from the results of other investigators that the flux of neutron-producing radiation varies exponentially with x and that the absorption mean free path has the value $L = 320$ g/cm². This will be confirmed by our own results. We therefore write

$$W(x) = W(0) \exp(-x/L), \quad (4)$$

where, from Eq. (1), $W(0) = R(0, 0) - B$. This assumption breaks down at small thicknesses, but at small thicknesses the last term in Eq. (3) is negligible in any case.

TABLE I. Counting rate resulting from neutrons produced in half-inch thickness of lead producer, as a function of absorber thickness.

Absorber thickness (inches)	Counting rate (counts/min) $R(x, p)$	$R(x, 0)$	$P(x)$
0	17.91 ± 0.14	11.36 ± 0.10	6.55 ± 0.17
0.25	19.13 ± 0.14	13.03 ± 0.11	6.21 ± 0.17
0.50	19.09 ± 0.16	14.08 ± 0.08	5.19 ± 0.18
1.0	20.17 ± 0.16	15.33 ± 0.13	5.12 ± 0.18
2.0	23.11 ± 0.19	18.81 ± 0.14	4.83 ± 0.24
4.0	28.20 ± 0.20	25.30 ± 0.19	3.89 ± 0.29

The results of the absorption experiment are summarized in Table I, where the indicated errors are standard deviations due to

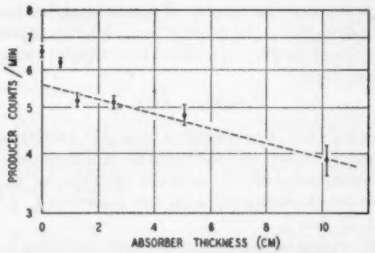


FIG. 2. Neutron production rate in half-inch thickness of lead placed under various thicknesses of lead absorber. The dotted line represents an exponential function with $L = 320$ g/cm².

counting statistics only. The function $P(x)$ is plotted in Fig. 2. The dotted line represents an exponential function with $L=320$ g/cm² and is seen to provide a good fit to the data for absorber thicknesses greater than ~1.2 cm. A transition maximum for zero absorber thickness is clearly indicated. Since $P(x)$ represents an integrated value for neutron production over a half-inch of lead, the transition maximum for neutron production per unit volume of lead must occur within the first 1.25 cm (half-inch) of lead.

* Assisted in part by the Flight Research Laboratory, USAF.
† AEC Predoctoral Fellow.
[1] E. P. George and A. C. Jason, Proc. Phys. Soc. (London) **A62**, 243 (1949).
[2] Bernardini, Cortini, and Manfredini, Phys. Rev. **76**, 1792 (1949); Phys. Rev. **79**, 952 (1950).
[3] Malaspini, Merlin, Pierucci, and Rostagni, Nuovo cimento **7**, 145 (1950).
[4] Schoffer, Höcker, and Kuhn, Phys. Rev. **82**, 444 (1951).
[5] Belliboni, Fabbrichesi, De Marco, and Merlin, Nuovo cimento **8**, 374 (1951).
[6] J. J. Lord and M. Schein, Phys. Rev. **75**, 1956 (1949).
[7] Shapiro, Stiller, Birnbaum, and O'Dell, Phys. Rev. **83**, 455 (1951).
[8] M. M. Shapiro and A. F. Gabrysh, Phys. Rev. **84**, 160 (1951).
[9] Heitler, Powell, and Heitler, Nature **146**, 65 (1940).
[10] J. A. Simpson (to be published).
[11] Cocconi, Tongiorgi, and Widgoff, Phys. Rev. **79**, 768 (1950).

Neutron Diffraction Investigation of the Atomic Magnetic Moment Orientation in the Antiferromagnetic Compound CrSb

A. I. Snow
University of Chicago, Chicago, Illinois
(Received November 23, 1951)

THE atomic magnetic moment orientation in the antiferromagnetic compound CrSb has been investigated at room temperature by neutron diffraction. The crystal structure and magnetic properties of this compound have been previously determined by Haraldsen, Rosenqvist, and Grønvold.[1] This compound has the nickel arsenide structure and shows a marked decrease in magnetic susceptibility as the temperature is lowered from the antiferromagnetic Curie point (450°C). The magnetic mass susceptibility falls from 10.97×10^{-6} at 450°C to 3.3×10^{-6} at 20°C. Lattice constants are $a_0=4.127$A and $c_0=5.451$A.

Calculation showed that all antiferromagnetic magnetic moment orientations which lead to a magnetic unit cell, one side of which is a multiple of the atomic (x-ray) unit cell, would cause the presence of strong new diffraction maxima at low angles in a neutron diffraction pattern. Such new maxima were not observed, although they were looked for carefully. Experimental data showed that the (101) maximum had twice the intensity of the (102) and (110) maxima, respectively. On the basis of nuclear scattering, with no magnetic scattering, these three maxima would be essentially equal in magnitude. This doubling in intensity of (101) relative to (102) and (110), respectively, is consistent with the atomic magnetic moments being aligned perpendicular to (001) planes, i.e., along the c axis, in such a manner that they are aligned in the same direction (ferromagnetically) in any one (001) plane but oppositely directed (antiferromagnetically) in adjacent (001) planes. Nearest neighbor chromium atoms (2.726A apart) are thus aligned antiferromagnetically whereas second nearest neighbors (4.127A apart) are aligned ferromagnetically. One unit cell is shown in Fig. 1. The possibility that the magnetic moments may be aligned parallel to the (001) planes may be eliminated since this arrangement would lead to an unobserved strong maximum at low angles.

On the basis of the relative intensity data, a conservative estimate for the average number of electrons per atom whose magnetic moments are aligned in the manner described above is 2.7 ± 0.2, if the assumption is made that the orbital angular momentum contribution is completely quenched and the magnetic amplitude form factor of manganese[2] is used in the calculation.

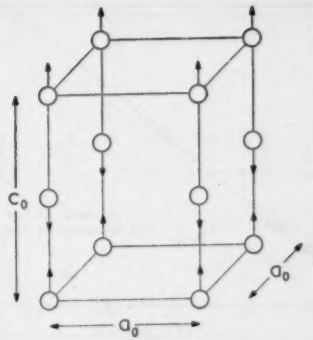


FIG. 1. Unit cell of CrSb. Arrows indicate the relative alignments of atomic magnetic moments. Only chromium atoms are shown.

The atomic magnetic moment orientation found above differs from that found in MnF_2 which is body-centered tetragonal with respect to the manganese atoms, where nearest neighbors (3.31A apart)[3] are ferromagnetically arranged and next nearest neighbors (3.82A apart)[3] are antiferromagnetically arranged.[4] The CrSb arrangement is somewhat similar to the arrangement in some of the (111) planes of the face-centered cubic compound FeO[2] since in FeO the moments are aligned perpendicular to some of the (111) planes and are aligned in opposite directions in adjacent (111) planes[4] of the particular sets considered.

Thanks are due to Dr. T. Rosenqvist for preparation of the compound and to Dr. G. R. Ringo and other members of the Argonne National Laboratory staff for their cooperation. The neutron diffraction work was performed by the author at the heavy water nuclear reactor of the Argonne National Laboratory.

[1] Haraldsen, Rosenqvist, and Grønvold, Arch. Math. Naturvidensk. B. L. Nr. 4.
[2] Shull, Strauser, and Wollan, Phys. Rev. **83**, 333 (1951).
[3] Calculated from lattice constants of M. Griffel and J. W. Stout, J. Am. Chem. Soc. **72**, 4351 (1950).
[4] C. G. Shull, Phys. Rev. **82**, 771(T) (1951).

Nutational Resonance*

H. C. Torrey
Department of Physics, Rutgers University, New Brunswick, New Jersey
(Received November 19, 1951)

APPLICATION of repeated radiofrequency pulses to proton-containing liquids in a magnetic field results in transient nuclear resonance absorption effects which we have reported[1] previously.

Recently we have observed an additional effect which results when both the time between pulses (off-time) and a pulse duration (on-time) are short compared with the relaxation times T_1 and T_2. Under these circumstances one might expect the nuclear resonance signal to be small in the case of power levels normally sufficient to saturate the sample. However, if a radiofrequency pulse contains an integral number of nutational periods it is possible for the signal to maintain a relatively large value as a result of a nutational resonance process. In case the magnetic field is homogeneous over the sample, the nutational frequency will be uniform and after an initial rf pulse lasting for an integral number of nutational periods the sample's magnetic moment will be left oriented along the magnetic field in its original condition except for some attenuation due to the net effect of the transverse and longitudinal relaxation processes. During the succeeding off-time, the moment's orientation will be undisturbed while its amplitude increases due to longitudinal relaxation. After a sufficient number

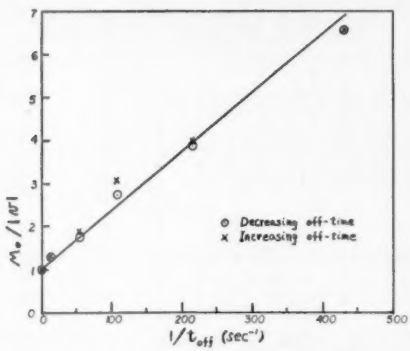


FIG. 1. Nutational resonance in distilled water.

of pulses a steady state results in which there is considerable signal strength even if the off-time is short compared with T_1 and T_2.

It is clear that this is a resonance process since, if the condition on the nutational frequency is not exactly fulfilled, the moment will make increasing angles with the magnetic field on succeeding pulses and in the steady state the signal will be small. It is not necessary, however, for the rf resonance condition $\omega=\gamma H_0$ to be exactly satisfied.

To explain the effect quantitatively we have solved the Bloch equations for the case of an infinite succession of radiofrequency pulses. This was done by first solving the Bloch equations for a single pulse with undetermined initial conditions and then applying the condition that the components of magnetic moment at the end of an off-time should be the same as at the start of an on-time. The general steady-state solution is somewhat involved, but the special case of radiofrequency resonance, $\omega=\gamma H_0$, yields for the amplitude

$$|v|/M_0=[1+(t_1/2t_2)(1+T_1/T_2)]^{-1} \times(1+A^2\sin^2\tfrac{1}{2}\theta)^{\frac{1}{2}}(1+AB\sin^2\tfrac{1}{2}\theta)^{-1},$$

where

$$A=2[\tfrac{1}{2}t_1(1/T_1+1/T_2)+t_2/T_2]^{-1}$$
$$B=2[\tfrac{1}{2}t_1(1/T_1+1/T_2)+t_2/T_1]^{-1}$$
$$\theta=\gamma H_1 t_1$$

and t_1 and t_2 are the on-time and off-time, respectively. It is assumed that both t_1 and t_2 are small compared with T_1 and T_2 so that A and B are very large. If then $\sin\frac{1}{2}\theta$ is not small we have

$$|v|/M_0\simeq\tfrac{1}{2}(t_2/T_1)|\csc\tfrac{1}{2}\theta|\ll 1.$$

If, however, $\theta=2n\pi$ where n is an integer (the nutational resonance condition), we get

$$|v|/M_0=[1+(t_1/2t_2)(1+T_1/T_2)]^{-1}, \qquad (1)$$

which may be of the order of unity.

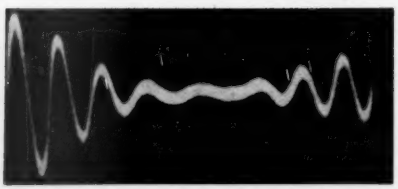


FIG. 2. Effect of an inhomogeneous field.

We have tested this result for distilled water with the result shown in Fig. 1, where we plot $M_0/|v|$ *versus* the reciprocal of the off-time holding t_1 fixed. The points fall well on a straight line in accord with Eq. (1). In this case the on-time contained one complete nutational period. From the slope of the line we obtain $T_2=0.63T_1$. Taking[2] $T_1=2.3$ sec we find $T_2=1.4$ sec. This is much the longest T_2 yet measured and confirms the conjecture[2] that the line width in distilled water is of the order of 10^{-4} oersted.

We believe that this phenomenon explains another effect first observed in this laboratory by Thomas.[3] If the magnetic field is sufficiently inhomogeneous the nutational frequency varies over the sample. The net signal is then the result of superposed sine waves from those parts of the sample obeying the nutational resonance condition. These sine waves will have different frequencies but all begin and end in the same phase. Thus there will be constructive interference at the start and at the end of a rf pulse but destructive interference in the center. A typical signal is shown in Fig. 2. This effect appears to be similar to one recently reported by Bradford, Clay, and Strick.[4]

* This work has been supported by the joint program of the ONR and AEC, by the Rutgers University Research Council, and by Radio Corporation of America.

[1] H. C. Torrey, Phys. Rev. **76**, 1059 (1949).
[2] Bloembergen, Purcell, and Pound, Phys. Rev. **73**, 679 (1948).
[3] J. T. Thomas, thesis, Rutgers (1950).
[4] Bradford, Clay, and Strick, Phys. Rev. **84**, 157 (1951).

A Solar Component of the Primary Cosmic Radiation*†

J. A. SIMPSON, W. FONGER, AND L. WILCOX
Institute for Nuclear Studies, University of Chicago, Chicago, Illinois
(Received November 26, 1951)

THE time-dependent changes of intensity of the low energy primary cosmic radiation as observed by changes in the nucleonic component intensity have been investigated further to determine the extent to which these changes may identify the origin of at least a small fraction of the low energy end of the primary particle intensity spectrum. Continuous observations are made at different geomagnetic latitudes and altitudes by measuring the disintegration product neutrons[1] from the local production of small stars in lead. The geometries are similar to those described earlier.[2] We report here preliminary results selected as typical samples of the behavior of the low energy primary radiation with time and restrict our present discussion to large world-wide changes which persist for the order of days.

The total number of events registered for successive twelve-hour intervals has been plotted for three of the monitor stations in Fig. 1. For example, each point for the Climax, Colorado, monitor (11,000 feet) represents more than 4×10^6 events. All monitors are normalized periodically with Ra-Be standard neutron sources. Hence, the principal error is the barometric correction. In spite of the differences in atmospheric temperature changes between Climax and Sacramento Peak, there appears a remarkable similarity of details of intensity changes among the three monitors.

By observing the change of intensity with latitude, λ, at 30,000 ft pressure altitude from $\lambda=40°$ to 63° N in aircraft on August 7, 18, and 25, 1951 we have been able to show the following:

a. The changes of intensity observed in Fig. 1 are of the type reported earlier,[1] and, hence, represent near the maxima in Fig. 1 additional primary radiation arriving at latitudes at least up to $\lambda=56°$ N.

b. Changes in the dipole magnetic field of the earth do not account for the observed changes of intensity.

Hence, these changes of cosmic-ray intensity are reasonably certain of being extra-terrestrial in origin.

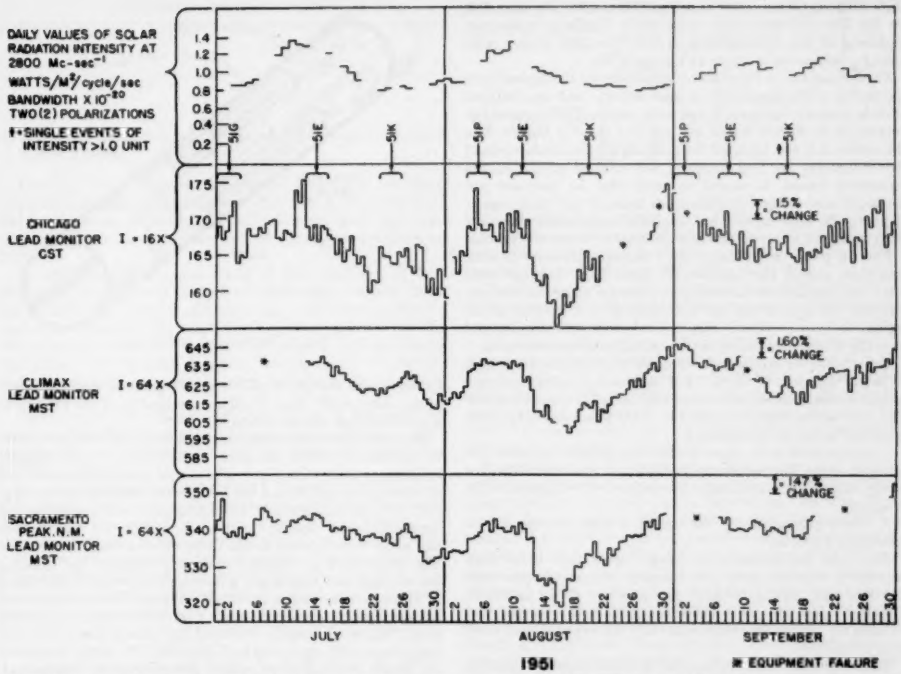


FIG. 1. The changes of neutron intensity, I, from the nucleonic component are shown as a function of time. The Chicago curve is for sea level observation, $\lambda = 52°$ N; the Climax curve is for 11,000 ft at $\lambda \approx 50°$ N and the lowest curve is for 9300 ft at $\lambda \approx 41°$ N. The solar radio noise at 2800 Mc/sec as measured by Covington (reference 3) is shown for comparison. The arrows indicate the central meridian positions for solar active regions 51*E*, 51*K*, 51*P*, and 51*G*.

A study of intensity data from 1950 and 1951 to date has shown a remarkable recurrence of series of maxima with each series having an approximately 27-day period. For example, in Fig. 1 the maximum located near 14 July has appeared for the preceding four 27-day periods and is again observed on 11 August. Another approximately 27-day sequence is observed with maxima near 24 July, 20 August, and 16 September.

By calling the date of any intensity maximum the nth day and each successive 27th day the nth day, a curve for 27-day periods has been plotted in Fig. 2 using the Chicago monitor data covering five 27-day periods, including data through 27 August but excluding data for June, 1951. (The justification for this will be shown in a later paper.) The maximum at n in Fig. 2 corresponds to the 27-day series maxima in Fig. 1 near 14 July and 11 August.

Since the sun has a 27-day spin at low latitudes, the maxima in Figs. 1 and 2 are found to correspond to local active regions on the sun. It will be shown in a later paper that a unique correlation between these maxima and the local active solar regions occurs for a coincidence within ± 1 day between the cosmic-ray maximum and the central meridian (c.m.) position of the active region. We define a local active region as an area producing intense coronal emission, flares, and enhanced radio noise[3] and bursts. By these criteria an active solar region was at c.m. on 14 July and 11 August, which corresponds to the intense corona region designated by the

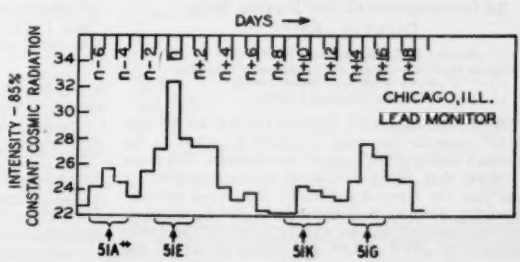


FIG. 2. Five 27-day cycles have been superposed for the Chicago station. The intensity peaks at $n-5$, n, $n+10$, and $n+15$ have been identified with local active regions on the sun at central meridian within ± 1 day. **Region 51A was replaced by 51P in August.

High Altitude Observatory[4] as region 51*E* crossing the solar disk for the fifth and sixth time, respectively. Similarly, a sequence beginning 24 July was identified as 51*K*. The 2800 Mc/sec radio noise for this period is shown at the top of Fig. 1.

Further validity of these data is obtained from the approximate correlation of the magnitude of solar activity and the observed particle intensity changes. A new solar region (51*P*) became important on its August transit and was at c.m. on 5 August. Also the region 51*E* was reported decaying on the September transit thus decreasing the importance of 51*E* relative to 51*P* on the September transit. It should be noted that we have not yet developed any method to distinguish between two local regions at nearly the same heliocentric longitude except when there is a great difference between the observed activity of the two regions.

From a group of 22 consecutive neutron intensity maxima coincident among the monitors, 19 have been identified with solar c.m. positions for active regions, two are uncertain, and one although real does not appear to correspond to a reported active region. It may be shown that the intensity decreases do not have any simple relationship to the onset of geomagnetic storms.

Tentatively, we call the solar component of the primary cosmic radiation that part of the low and intermediate energy primary radiation which shows extra-terrestrial intensity changes associated with active regions on the sun. This solar component may arise from either of the following:

1. Redistribution of some cosmic-ray particle energies (for example, when the cosmic-ray particles are accelerated by low energy neutral ion beams emitted from local active regions on the sun);[5] or

2. Direct production by the sun of a small fraction of the cosmic-ray particles.[6]

Since the measurements at $\lambda \approx 41°$ show that these time-dependent changes occur for primary proton energies even >5 Bev and there is no large diurnal effect at the maximum intensity, there are some difficulties with both concepts, but perhaps the greater difficulties exist with the assumption of direct solar production.

These experiments will be reported in detail covering the period 1950 and 1951. The authors wish to thank Mr. S. B. Treiman for assistance in the early stages of the work.

* Assisted in part by the Flight Research Laboratory, USAF.
† Bull. Am. Phys. Soc. **26**, No. 6, 24 (1951).
[1] J. A. Simpson, Phys. Rev. **81**, 893 (1951); Phys. Rev. **83**, 1175 (1951), Sec. VI.
[2] J. A. Simpson, Proc. Echo Lake Conf., p. 252 (1949).
[3] 2800 Mc/sec data were kindly furnished by Mr. A. Covington, National Research Council, Ottawa, Canada. 200 Mc/sec data are kindly provided by Dr. C. Burrows and the Radio Astronomy Group, Dept. of Electrical Engineering, Cornell University, Ithaca, New York.
[4] These corona data were prepared by Dr. W. O. Roberts and Miss Dorothy Trotter, High Altitude Observatory, Boulder, Colorado. We are indebted to them for the careful preparation of isophotal charts of the 5303A emission and their close cooperation in reporting solar flare effects. The assistance of Dr. A. H. Shapley and the Central Radio Propagation Laboratory, National Bureau of Standards, in providing prompt S. I. D. and flare data is greatly appreciated.
[5] H. Alfvén, Phys. Rev. **75**, 1732 (1949).
[6] See, for example, E. M. McMillan, Phys. Rev. **79**, 498 (1950).

An Investigation of the Double Beta-Decay of $_{50}Sn^{124}$ *

MARVIN I. KALKSTEIN† AND W. F. LIBBY
Institute for Nuclear Studies and Department of Chemistry, University of Chicago, Chicago, Illinois
(Received December 3, 1951)

FIREMAN[1] has reported a half-life between 4×10^{15} and 9×10^{15} years for the decay of $_{50}Sn^{124}$ to $_{52}Te^{124}$ by the simultaneous emission of two negative beta-particles. This result would indicate that double beta-decay is unaccompanied by neutrinos, since the expected half-life for the process with the emission of two neutrinos is of the order of 10^{19} times longer. However, the results of Inghram and Reynolds[2] for the double beta-transition of $_{52}Te^{130}\rightarrow_{54}Xe^{130}$ and Levine *et al.*[3] for the double beta-transition of $_{92}U^{238}\rightarrow_{94}Pu^{238}$ give minimum half-lives of the order of 10^{19} years. Further investigation of the activity of $_{50}Sn^{124}$ therefore seemed desirable.

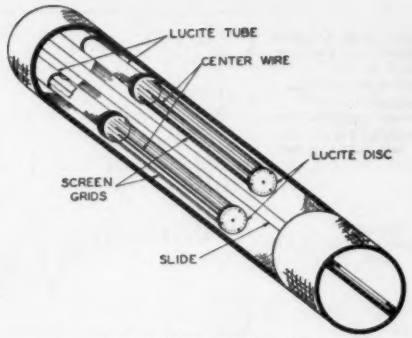


FIG. 1. Double screen wall counter.

Our method of measurement involves two screen wall counters[4,5] in coincidence within the same 24-inch long, 3-inch diameter cylinder, as is shown in Fig. 1. The sample is mounted on one-half of a 16-inch long slide and the background material on the other half. Alternate counting of the background and sample can be accomplished by tipping the counter, so that the sample slide can move from one end to the other inside a fixed center slide. The total counting volume, for the two halves of the cylinder, has an eight-inch length and a three-inch diameter. This enables us to measure samples of up to 140 cm² area. The two screen wall counters are in coincidence, while a set of 11 counters, 18 inches long, 2 inches in diameter, surround the cylinder and are in anti-coincidence with the screen wall counters. The entire setup is in an 8-inch iron shield to reduce the gamma-ray background. Tests showed that coincidences between the two halves occur at a rate expected if the counter were sensitive all the way to the corners of the hemicylindrical volumes defined by the counter cylinder and the sample foil.

A 3.013-gram sample of stannic oxide containing 95±1 percent Sn^{124} was obtained from the Stable Isotopes Research and Production Division of the Oak Ridge National Laboratory. The sample was reduced to the metal by mixing it with a small excess of sodium cyanide and heating for several hours. This was then rolled to a 24-mg/cm² thickness and 83-cm² area. Some pure stannic oxide, of ordinary isotopic abundance, was treated in exactly the same fashion and used as the background material. Since investigations of the coincidence background showed a marked dependence on the thickness of slide material due to scattering of chance gamma-rays, the background metal was rolled to the same thickness and then cut to the same shape and area as the sample.

Both the singles and coincidence activity of both the sample and background were measured and the results are listed in Table I. The errors given are standard deviations calculated from the counting statistics.

TABLE I. Measured activities of enriched and ordinary tin.

Activity measured	Background (cpm)	Sample (cpm)
Singles (half-cylinder) without A.C.'s*	90.5 ±1.1	89.3 ±1.0
Singles (half-cylinder) with A.C.'s	8.40 ±0.15	8.23 ±0.16
Coincidence without A.C.'s	38.1 ±0.2	38.2 ±0.2
Coincidence with A.C.'s	1.310±0.007	1.304±0.007

* Anticoincidence shielding.

TABLE II. Minimum half-lives for various energies from measurement of enriched tin sample.

Energy (Mev)	Minimum half-life (years)
1.0	1.7×10^{17}
1.5	2.1×10^{17}
2.0	2.4×10^{17}

Assuming that the maximum possible activity is less than twice the standard deviation, we can calculate a minimum half-life for the double beta-decay of Sn^{124}. The minimum half-lives as a function of the energy assumed for the process are given in Table II.

An investigation also was made of the possibility of two successive beta-emissions through an intermediate nucleus. Antimony carrier was added to nine moles of pre-war stannic chloride and then separated. The antimony was then counted for the 60-day beta-particle of Sb^{124}. An activity of 0.38 ± 0.23 cpm above background was obtained. Assuming the activity for the sample to be less than 0.5 cpm, the minimum half-life for the transition is 10^{17} years.

These results differ significantly from those of Fireman.[1] In order to obtain his result, we would have had to observe at least 0.5 cpm due to the sample, whereas our measured activity was -0.006 ± 0.010 cpm.

A number of possible explanations for Fireman's activity, other than double beta-decay, suggested themselves during the course of this work. One possibility is that his two samples were not of the same thickness. If there were any gamma-ray contamination in the samples or elsewhere in the experimental environment and if his enriched sample was thinner than the other, a higher coincidence rate for the enriched sample would result. Another possibility is contamination of his enriched sample with a beta-activity that would give coincidence counts by backscattering from one counter into the other.

Appreciation is expressed to Dr. C. P. Keim of the Stable Isotopes Research and Production Division of the Oak Ridge National Laboratory for his cooperation in making available the enriched Sn^{124} isotope used in this work. The experiment on the removal of antimony from a large tin sample and testing it for Sb^{124} was proposed by Dr. T. P. Kohman.

* This research was supported in part by the Air Force, Wright Air Development Center.

† AEC Predoctoral Fellow. Now at Radiation Laboratory, University of California, Berkeley, California.

[1] E. L. Fireman, Phys. Rev. **75**, 323 (1949).
[2] M. G. Inghram and J. H. Reynolds, Phys. Rev. **76**, 1265 (1949).
[3] Levine, Ghiorso, and Seaborg, Phys. Rev. **77**, 296 (1950).
[4] W. F. Libby, Phys. Rev. **46**, 196 (1934).
[5] W. F. Libby and D. D. Lee, Phys. Rev. **55**, 245 (1939).

L-Electron Capture and Alpha-Decay in Np^{235}*

RALPH A. JAMES,† ALBERT GHIORSO, AND DONALD ORTH‡
Radiation Laboratory, University of California, Berkeley, California
(Received December 4, 1951)

THE isotope Np^{235} has previously been reported[1] to decay by orbital electron capture with a half-life of approximately 400 days. By the bombardment of 10 grams of enriched U^{235} (95 percent U^{235}) with 20-Mev deuterons in the 60-inch Crocker Laboratory cyclotron, a much larger sample has been prepared and more accurate measurements made on the radiations of Np^{235}.

After allowing the bombarded target to decay for several months so that all short-lived activities had completely decayed, a very pure neptunium fraction was isolated. This sample, in addition to the x-rays accompanying electron capture, emitted a small number of alpha-particles having an energy of 5.06 ± 0.02 Mev (disintegration energy 5.15 Mev). The decay of these alpha-particles and the x-rays have now been followed for a little more than 2 years and both give a half-life of 410 ± 10 days, showing that both originate from decay of Np^{235}.

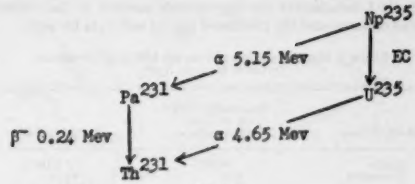


FIG. 1.

The disintegration energy for the reaction $Np^{235}+e^-=U^{235}$ can be calculated using the closed cycle system and the disintegration energies shown in Fig. 1. These values give $5.15-(4.65+0.24)=0.26$ Mev as the energy avilable for capture of a free electron. After subtracting the orbital electron binding energies, 0.14 Mev and 0.24 Mev are the energies available for *K*-electron capture and *L*-electron capture, respectively. This very low disintegration energy and the high nuclear charge are expected[2] to give a high ratio of *L*-electron captures to *K*-electron captures, particularly in the case of forbidden transitions.

In agreement with this prediction, a large excess of *L* x-rays over the number of *K* x-rays has been observed, corresponding to greater than 90 percent *L*-electron capture and less than 10 percent *K*-electron capture. This conclusion was first reached from absorption curves using the best available estimates of the counting efficiencies of these *K* and *L* x-rays and has since been confirmed by the use of xenon filled proportional counters connected through a linear amplifier to a 48-channel pulse analyzer in an arrangement similar to that of Hanna *et al.*[3] This last method also gives evidence that most of the *L* x-rays arise from vacancies in the L_{II} and L_{III} levels. These vacancies can be created either by (1) capture of electrons from the L_{II} and L_{III} levels or (2) capture of an electron from the L_I level followed by a transition of the Coster-Kronig type to the L_{III} level. The available data do not seem to warrant an estimate of the relative numbers of captures from the various *L* levels but do warrant the conclusion that at least 90 percent of the electrons captured are from the *L* levels.

Figure 2 shows a typical curve obtained on the sample with settings of counter tube voltage and grid bias voltage corresponding to an energy range of about 7–27 kev. The observed curve (solid line) has the general shape expected from the *L* x-rays arising from the L_{II} and L_{III} levels while the dotted line shows the general shape expected if they originated in the L_I level.

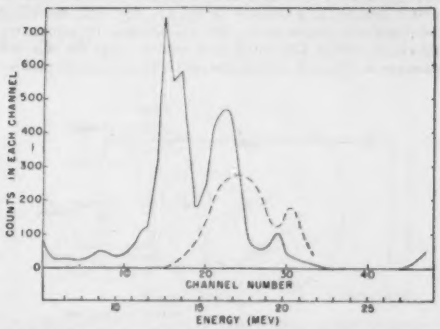


FIG. 2. Pulse-analysis curve of *L* x-rays emitted by Np^{235}. Dashed curve is general form of curve expected for transitions from the L_I level.

Table I summarizes the approximate amount of the various modes of decay and the calculated partial half-lives for each.

TABLE I. Branching ratios and partial half-lives for various modes of decay in Np^{235}.

Mode of decay	Branching ratio (% of total disintegrations)	Partial half-life
Alpha	0.005	2×10^6 yr
K-capture	<10	>12 yr
L-capture	>90	<450 days
Total	100	410 days

The authors are indebted to Professor Glenn T. Seaborg under whose direction this problem was undertaken, and to G. B. Rossi and the 60-inch cyclotron crew for the excellent U^{235} bombardment.

* This work was performed under the auspices of the AEC.
† Present address: Department of Chemistry, University of California, Los Angeles, California.
‡ Present address: Oak Ridge National Laboratory, Oak Ridge, Tennessee.

[1] James, Florin, Hopkins, and Ghiorso, *The Transuranium Elements: Research Papers* (McGraw-Hill Book Company, Inc., New York, 1949), Paper No. 22.8, National Nuclear Energy Series, Plutonium Project Record, Vol. 14B, Div. IV.
[2] R. E. Marshak, Phys. Rev. **61**, 431 (1942).
[3] Hanna, Kirkwood, and Pontecorvo, Phys. Rev. **75**, 985 (1949).

Charged π-Meson Production from Deuterium*

S. PASSMAN, M. M. BLOCK,† AND W. W. HAVENS, JR.
Columbia University, New York, New York
(Received November 30, 1951)

EXPERIMENTAL RESULTS.—The $\pi^{\pm}$ meson production cross sections from deuterium at 90°±5° to a 381-Mev proton beam have been obtained. The technique utilizing nuclear emulsions embedded in a tapered copper absorber, exposed to the internal circulating beam of the Nevis cyclotron, has been described previously.[1] The experimental arrangement has been modified to obtain better energy resolution. The target is now placed radially, and a copper "clipper," to define the beam energy, is placed ¼ inch radially behind the leading edge of the target, and located 90° from the target in azimuth, to prevent neutrons scattered from the "clipper" from striking the emulsions (see Fig. 1).

The target, 60 mils thick, was made from deuterated paraffin, whose composition was analyzed to be ~96 percent $(CD_2)_n$.[2] It was exposed for 7 minutes to an average proton beam current of $\sim10^{-10}$ ampere, at an energy of 381±5 Mev. The deuterium contribution to meson production was obtained by subtracting the known carbon differential cross section[1] from the observed spectrum of CD_2, and halving this result. The deuterium positive

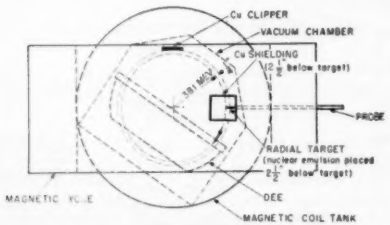


FIG. 1. Floor plan of Nevis cyclotron, showing experimental arrangement for proton bombardment of radial target.

meson spectrum, $d^2\sigma/d\omega dE$, obtained from $174\pi-\mu$ decays is shown in Fig. 2. It has a maximum in the region between 30 and 50 Mev of about 4.5×10^{-30} cm² Mev⁻¹ sterad⁻¹ D-nucleus⁻¹. The integrated cross section $d\sigma^+/d\omega$ is $(2.9\pm1.2)\times10^{-28}$ cm² sterad⁻¹ D-nucleus⁻¹. In the same regions scanned for the positives, only $6\pi^-$ meson stars were observed from $(CD_2)_n$. The statistical errors are too large to permit the drawing of a spectrum but an estimate of the total 90° negative cross section, $d\sigma^-/d\omega$, is $(1.1\pm0.9)\times10^{-28}$ cm² sterad⁻¹ D-nucleus⁻¹.

Analysis.—In analyzing the production of mesons in the bombardment of deuterium by highly energetic protons, we have used the impulse approximation;[3] that is, we have assumed that π-mesons are produced by collisions between an impinging proton and a single nucleon in the loosely bound deuteron. The second nucleon in the deuteron serves only to impart a deuteron momentum distribution to the struck nucleon prior to the collision, and satisfies conservation laws by carrying off the complimentary momentum.

The possible modes of producing mesons in nucleon-nucleon collisions between the incident proton and the proton or neutron of the deuteron are: (a) $p+p\rightarrow\pi^++d$ (or $p+n$); (b) $p+n\rightarrow\pi^+ +n+n$; (c) $p+n\rightarrow\pi^-+p+p$.

Although there is no direct experimental evidence on the relative magnitudes of processes (b) and (c), these reactions probably have comparable cross sections. This is due to charge independence of the low energy nuclear forces involved in the final state nucleon

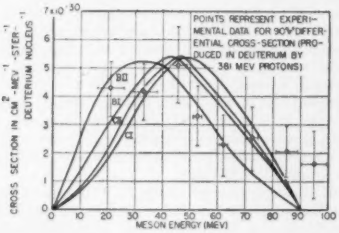


FIG. 2. Theoretical curves for the π^+ meson spectra from deuterium, labeled to indicate the energy and angular dependence used in the calculation.

interactions of these reactions, as well as the symmetrical properties of the π^+ and π^- mesons, i.e., lifetime, mass, charge, etc. The experimental (π^+/π^-) production ratio in deuterium, measured at 90° to a 381-Mev proton beam, is very large, ~25.[4] This indicates the relative smallness of process (c) [and in turn from the above discussion, the smallness of process (b)] relative to process (a). The strong interaction of nucleons in the final state of process (a) may account for the relative importance of this mode of meson production. The small cross section for process (c) may also be the effect of the exclusion principle, since there are 3 low energy protons present in the final state when the remaining proton in the deuteron is taken into account.[5]

With the hypothesis that process (a) is the main reaction contributing to π^+ meson production in deuterium, a theoretical calculation of this latter meson spectrum has been made based upon these assumptions:

(1) The interaction of the colliding nucleons in the final state is sufficiently large to give the meson the maximum available energy. From this it follows that the density of states is proportional to $\sqrt{T_M}$. Here T_M is the maximum energy available to the created meson in the center-of-mass system of the two colliding nucleons.

(2) The matrix elements for meson production in $p-p$ collisions are energy dependent, being proportional to $\sqrt{T_M}$ (type B), or to T_M (type C).

Assumption 2 is based upon analysis of the excitation function for meson production by protons in hydrogen.[1] Assumption 1 is consistent with the experimental evidence obtained from the spectral shape of mesons produced by protons in hydrogen.[6] It is also similar to the hypothesis used by the authors in interpreting the heavy element meson production spectra.[1]

The calculated 90° π^+ meson spectrum in deuterium was obtained by using the $p-p$ cross sections implied in assumptions (1) and (2) above, and summing over the deuterium momentum distribution derived from a Fourier transform of the deuteron ground-state wave function of the form $e^{-\alpha r}/r$. The calculations took into consideration the recoil of the third nucleon present in the collision. Two types of angular distribution were considered for the created mesons in the c.m. system, isotropic (type I) and $\cos^2\theta$ (type II).

The theoretical curves for the π^+ meson spectra from deuterium, labeled to indicate the energy and angular dependence used in each calculation and normalized to give the observed 90° cross section, are plotted in Fig. 2. Within the limited statistical accuracy of the data, the calculations are in agreement with experiment with regard to the shape and position of the maximum of the spectrum.

These calculations indicate that the proton in the deuteron has a cross section 2.2 (*CI*), 1.7 (*CII*), 1.6 (*BI*), 1.2 (*BII*) times that of a free proton for π^+ production by protons at 381 Mev. The experimentally determined 90° π^+ cross section from deuterium is (5.6±2.3) times that of the free proton cross section at 381 Mev. Therefore, unless further experiments, now in progress, show this ratio to be near the lower statistical limit given above, it is clear that it will require an excitation function of meson production greater than the $\sqrt{T_M}$ or T_M matrix element dependence used in the above calculations, or thought necessary from experiments on meson production in hydrogen, to explain the large π^+ production ratio of deuterium to hydrogen. Another possibility is that the assumption of an impulse approximation in treating the two nucleons in deuterium as independent particles for the purpose of meson production may be in error; perhaps the interaction of the proton with the deuteron as a whole must be considered. However, in this latter possibility the reaction $p+D\rightarrow T+\pi^+$ might be appreciable, with the resulting mesons coming off in a line spectrum at about 85 Mev in our experiment. No such "tritium" meson peak at the high energy end of the spectrum was observed.

* This work was assisted by the joint program of the ONR and AEC.
† Now at Duke University, Durham, North Carolina.
[1] Block, Passman, and Havens, Phys. Rev. **83**, 167, 205 (1951).
[2] The deuterated paraffin was prepared by the Texas Oil Company, which gave a composition of (97.2±2) percent deuterium. A spectroscopic analysis of the sample used in the experiment was made by Professor T. I. Taylor of the Columbia University Chemistry Department, and he found a composition of (96±2) percent deuterium.
[3] G. Chew, Phys. Rev. **80**, 196 (1950).
[4] Professor G. Bernardini has kindly informed us that his latest experiments indicate that π^- mesons interact more strongly with matter than do π^+ mesons of the same energy. Since this experiment involves a slowing down of the π^+ mesons in Cu absorbers, this asymmetry in nuclear interactions must be taken into account. However, the total correction factor for a geometrical nuclear absorption cross section, as previously applied, amounts to < a 25 percent increment to the total production cross section, so that the large (π^+/π^-) production ratio from D reported above is not significantly altered by the quoted results.
[5] In a more detailed theoretical analysis of the production of mesons in deuterium, H. P. Noyes [Phys. Rev. **81**, 924 (1951)] indicates that the (π^+/π^-) ratio at 0° to a 345-Mev proton beam could be as large as 8.2 because of the above-mentioned effects.
[6] F. Cartwright, Ph.D. thesis, University of California Radiation Laboratory Report No. 1278 (1951).

Three-Quantum Annihilation and Positronium*

S. DeBENEDETTI AND R. SIEGEL†
Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received November 28, 1951)

THE three-quantum annihilation of positrons and electrons in the triplet state has been studied theoretically by Ore and Powell,[1] among others, and experimentally by Rich;[2] the work of Deutsch on positronium[3] is closely related to this effect, and can be considered an implicit proof of its existence. With the present experiment we have directly detected the three γ-rays of annihilation by means of a triple coincidence method.

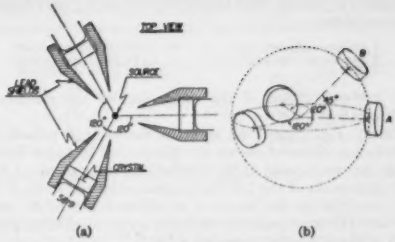


FIG. 1. Arrangement of the three counters for the detection of three-quantum annihilation.

The apparatus is shown in Fig. 1. The source used was Na^{22}. The three scintillation counters (5819 photomultipliers and NaI(Tl) crystals 4 cm in diameter and 2.5 cm thick) subtended solid angles $\Omega/4\pi=8\times10^{-3}$. They were connected to differential pulse-height selectors (*DPHS*) and to a first coincidence circuit of 10^{-7} sec resolving time. The outputs from the three *DPHS*'s and from the first coincidence circuit were connected to another quadruple coincidence circuit of $\sim2\times10^{-8}$ sec resolving time. The experiment consisted of measurements of coincidence rate with the counters coplanar with the source [*A*, Fig. 1(b)] and measurements of background with one of the counters rotated 45° out of this plane [*B*, Fig. 1(b)]. The bands accepted by the pulse height selectors were varied for a study of the energy of the coincident rays.

In a first experiment the positrons from a source of about 10^6 disintegrations/sec were allowed to stop in a solid. With the *DPHS*'s set to accept electron pulses of energy between 150 and 500 kev, the following counting rates were observed:

Counters coplanar	2.42±0.14 counts/min
Counters not coplanar	1.37±0.14 counts/min
Difference	1.05±0.19 counts/min

It seems difficult to account for this difference in any other way than by three-quantum annihilation. The magnitude of the effect agrees within a factor 2 with the theoretical expectation of one triplet annihilation per 370 singlet annihilations; however, our knowledge of counter efficiency and source strength is at present too poor for a more quantitative comparison.

In order to verify the formation of positronium in freon, as shown by the beautiful experiments of Deutsch, a source of about the same strength was deposited on thin Al and placed at the center of a bell-shaped Al container 1.2 cm in radius. The container was filled with freon 12 (dichlorodifluoromethane) at six

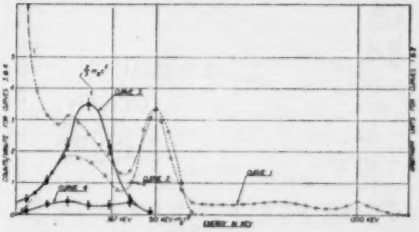


FIG. 2. Pulse-height distributions: (1) single counts from Na^{22}; (2) coincident pulses from two-quantum annihilation; (3) coincident pulses from three-quantum annihilation; (4) background for three-quantum annihilation.

atmospheres pressure. With this source the counting rates observed were as follows:

Counters coplanar	11.85 ±0.43 counts/min
Counters not coplanar	1.33 ±0.15 counts/min
Difference	10.52 ±0.46 counts/min

The tenfold increase of the effect confirms the stability of positronium in freon.

Finally, Fig. 2 shows the results of the study of pulse-height distribution. Curve 1 shows the distribution of pulses from a single counter exposed to the γ-radiation from a source of Na^{22}. The photopeaks of the 1.3-Mev nuclear γ-ray and of the 0.510-Mev annihilation line are used to calibrate the *DPHS* scale; another calibration point is provided by the position of the photopeak of the 0.367-Mev γ-rays from I^{131}.

Curve 2 was obtained with a double coincidence arrangement responding to two-quantum annihilation; it shows the pulse-height distribution of the coincident pulses from one of the two counters. It is evident that only the effect of the mc^2 γ-rays is present here. Curve 3 was obtained with the triple coincidence setup of Fig. 1 responding to the triplet annihilation of the source in freon, and shows the distribution of the coincident pulses in any one of the three counters. The single peak at $\frac{2}{3}$ mc^2 is further evidence for the three-quantum effect. Curve 4 is the corresponding background measured as described above.

Further studies of the effect in solids and gases are in progress.

* Supported by the AEC.
† AEC Predoctoral Fellow.
[1] A. Ore and J. L. Powell, Phys. Rev. **75**, 1696 (1949).
[2] J. A. Rich, Phys. Rev. **81**, 140 (1951).
[3] M. Deutsch, Phys. Rev. **82**, 455 (1951); Phys. Rev. **83**, 866 (1951); M. Deutsch and E. Dulit, Phys. Rev. **84**, 601 (1951).

Negative Ion Formation in Oxygen*

RONALD GEBALLE AND MELVIN A. HARRISON
Department of Physics, University of Washington, Seattle, Washington
(Received November 15, 1951)

WE have made measurements of ionization currents in O_2 which clarify certain points in connection with the Townsend coefficient and the attachment cross section in this gas. These measurements used plane-parallel electrodes separated as far as 40 mm and pressures from 11 to 40 mm of Hg in a range of E/p from 27.5 to 75 volts/cm/mm. Plotting, in the conventional way, log i against electrode separation, we find at each pressure a family of curves markedly different from the ordinary. Although the initial slope of such graphs has been reported as strictly linear for all gases investigated previously[1,2] we find O_2 exhibits an initial curvature followed by a more or less linear portion. The initial curvature becomes less prominent as E/p increases and is imperceptible for $E/p > 65$ at these pressures. We have observed similar curves in another gas, CF_3SF_5,[3] but will confine this discussion to the case of O_2. To check our apparatus we made careful measurements in H_2 but found no trace of a departure from linearity.

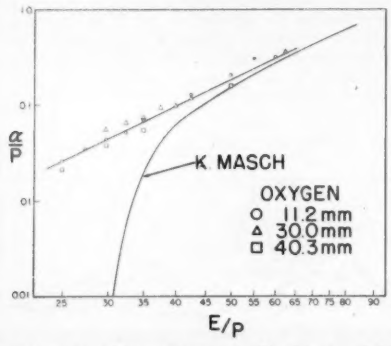


FIG. 1. Values of ionization coefficient α/p vs E/p in O_2 given by Masch (see reference 4) compared with those of the present experiment.

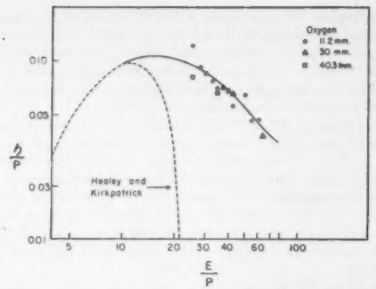


FIG. 2. Values of attachment probability η/p in O_2 given by Healey and Kirkpatrick (see reference 6) compared with those of the present experiment.

If ionization occurs in the gas at the mean rate α per cm and attachment forms negative ions (O_2^- and/or O^-) at the mean rate η per cm, the steady-state current is given by

$$i = i_0[\alpha \exp(\alpha-\eta)\delta - \eta]/[\alpha-\eta],$$

where δ is the plate separation and i_0 is the photoelectric current liberated at the cathode by an external source. This equation can be fitted closely to our experimental points. The process of curve-fitting on a basis of the foregoing assumptions yields values of both α/p and η/p.

The only published data with which to compare the present values of α/p are those of Masch[4] reproduced in Fig. 1. Masch attributed the rapid drop below $E/p \sim 40$ to attachment but did not attempt to analyze its effect. It can be seen from Fig. 1 that if due allowance is made for attachment in computing α/p, this drop, which does not appear in other gases, is not found in O_2.

Attachment probabilities in O_2 have been measured by Bradbury[5] and Healey and Kirkpatrick.[6] The curves submitted by these authors are similar in shape but differ by a factor of about 2 in magnitude. It is likely that this discrepancy arises only from differences in the values of drift velocity used to compute probabilities from the directly measured values of η/p. Figure 2 shows the values of η/p quoted by Healey and Kirkpatrick, which therefore are probably close to those (not quoted directly) found

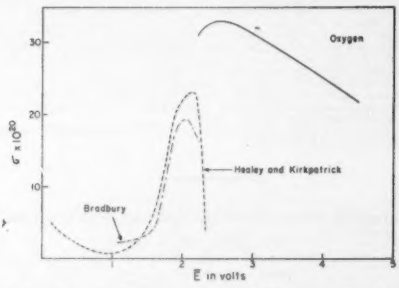


FIG. 3. Attachment cross section vs mean electron energy. Dashed curve—Healey and Kirkpatrick; dot-dashed curve—Bradbury; solid curve—present results computed from our η/p using thermal and drift velocities of Healey and Kirkpatrick.

by the former. It can be seen that instead of the steeply falling curve of the aforementioned authors, present data indicate a coefficient of appreciable magnitude at relatively large E/p. This circumstance can be explained by reference to our data on α/p; it is apparent that some multiplication is taking place. If proper allowance for this process is not made, the apparent attachment will be less than the true rate. It thus appears that the quantities α and η cannot be treated independently in regions where they are of comparable magnitude.

Figure 3 shows the attachment cross section plotted against mean electron energy computed from our measurements with the aid of the data of Healey and Kirkpatrick.[6] Previous explanations[6,7] of the maximum at 2 volts have been influenced by its apparent sharpness and have led to the conclusion that the negative ions were molecular. It now seems plausible that the reaction

$$O_2+e \rightarrow O^-+O$$

found by Lozier[8] and Hagstrum[9] with an appearance potential of 3.0 volts and a maximum at 7 volts is responsible for much of the high energy negative ion production. It is still possible that the process postulated by Bradbury[4,5] contributes to the steep rise that occurs at about 1.6 volts.

Further investigations of attachment processes in O_2 and other gases are in progress.

* This work was presented at the Conference on Gaseous Electronics, Schenectady, October 4–6, 1951.

[1] J. S. Townsend, *Electricity in Gases* (Oxford University Press, London, 1914).
[2] L. B. Loeb, *Fundamental Processes of Electrical Discharge in Gases* (John Wiley and Sons, Inc., New York, 1939).
[3] M. A. Harrison and R. Geballe, Phys. Rev. **83**, 884 (1951).
[4] K. Masch, Arch. Electrotech. **26**, 587 (1932).
[5] N. Bradbury, Phys. Rev. **44**, 883 (1933).
[6] R. H. Healey and J. W. Reed, *The Behavior of Slow Electrons in Gases* (Amalgamated Wireless, Sydney, 1941), p. 94.
[7] D. R. Bates and H. S. W. Massey, Phil. Trans. **A239**, 269 (1943).
[8] W. W. Lozier, Phys. Rev. **46**, 268 (1934).
[9] H. D. Hagstrum, Revs. Modern Phys. **23**, 185 (1951).

The Yield of Neutral Mesons from Proton Bombardment of Light Nuclei*

R. W. Hales, R. H. Hildebrand,† N. Knable,‡ and B. J. Moyer
Radiation Laboratory, Department of Physics, University of California, Berkeley, California

(Received December 3, 1951)

THIS experiment was planned in order to study the relative cross sections of various light nuclei for the production of neutral pions by 340-Mev protons.

In particular we wished to measure the yield of π^0's in $p-p$ collisions in order to obtain information about the symmetry properties of the π^0-meson.

If the π^0 is a pseudoscalar meson it cannot be emitted in a P angular momentum state from a $p-p$ collision in an even orbital state, since parity and angular momentum cannot be conserved simultaneously in this case. If the π^0 is emitted in an S state, then, in analogy with π^+-production for which the S state yield is about one-eighth of the P state yield, we should expect the numbers of π^0's from $p-p$ collisions to be about one-eighth of that from $p-n$ collisions. If the π^0 is a scalar meson no such large ratio would be expected.

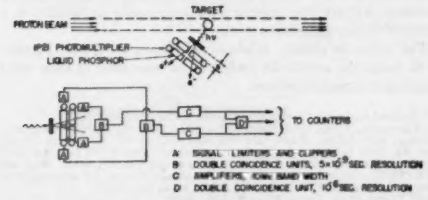


FIG. 1. Apparatus and circuit for detection of high energy photons.

TABLE I. Relative cross sections for neutral meson production by 341-Mev protons. The carbon yield has arbitrarily been given the value 6.00.

Nucleus	Relative cross section
H^1	0.07 ±0.06
H^2	0.85 ±0.20
Be^9	5.01 ±0.12
B^{10}	4.99 ±0.11
B^{11}	5.95 ±0.13
C^{12}	6.00 ±0.16
O^{16}	7.46 ±0.28

No direct comparison of the $p-p$ and $p-n$ π^0-yields could be made using a proton beam, but relative $p-p$ and $p-d$ yields were obtained by subtractions utilizing targets of carbon, polyethylene, heavy water, ordinary water and liquid oxygen. The series Be^9, B^{10}, B^{11}, C^{12} was also studied.

A unique method of identifying a neutral pion would be to detect both of the decay γ-rays in coincidence in the manner used for detection of neutral photomesons,[1] but the unfavorable neutron background conditions associated with proton bombarded targets makes the method infeasible for this experiment. Hence it was necessary to depend on single γ-detection. This is probably a valid measure of the neutral meson yield since the photon spectrum from proton-bombarded targets, which has been measured by Bjorklund, Crandall, Moyer, and York[2] and by Crandall and Panofsky,[3] is nearly consistent with a pure π^0-decay spectrum. The possible reabsorption of a charged or neutral meson within the nucleus in which it is created seems to be not more than a few percent.

The pair spectrometer method of Bjorklund *et al.*[2] was not suitable for comparing yields from different targets because of the uncertainties involved in multiple traversals when viewing internal cyclotron targets.

The photon detector used for this experiment was of the type developed by Hildebrand, Knable, and Leith,[4] in which multiple scattering in a thick converter is used to separate the electron pairs. In this device two scintillation counter telescopes are

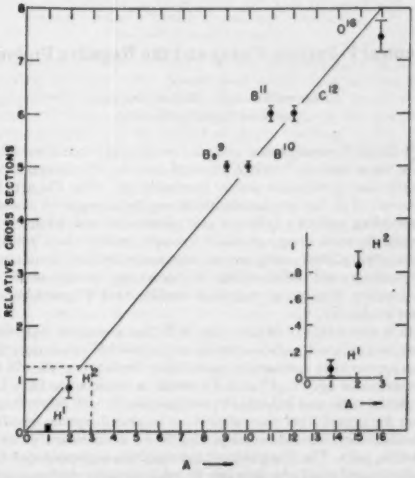


FIG. 2. Relative neutral meson yields.

placed behind the converter and connected in coincidence as shown in Fig. 1, so that whenever the electron and positron are scattered into separate telescopes the event is recorded. The phosphors were thick, so that only photons above 20 Mev could produce pairs with sufficient energy to cause a coincidence of all four counters. This detector has good discrimination against neutrons and charged particles and was thus suitable for use with the external beam in the presence of a high background.

The detector was placed at an angle of 135° to the direction of the proton beam for all these measurements in order to avoid the large flux of neutrons from the target which is present at small angles. The results obtained at this one angle should be a fairly good measure of relative total yields of π^0s, since the angular distribution of photons in the meson's rest frame is isotropic so that the photon yield at one angle in the lab frame receives contributions from mesons emitted in any direction. That the photon yield in a given direction in the laboratory is not strongly affected by conceivable meson angular distributions can be shown by straightforward calculations.

The results are shown in Table I and Fig. 2. The yields shown are relative to the carbon yield which has arbitrarily been given the value 6.00.

The yield from H^1 appears to be suppressed, in accordance with the selection rule discussed above for a pseudoscalar meson. Further measurements using a liquid hydrogen target in place of the carbon and polyethylene targets will be necessary to establish a positive yield from hydrogen or to lower the present upper limit appreciably. The yields from the series Be^9, B^{10}, B^{11}, C^{12}, which differ by the successive addition of a proton, a neutron, and a proton, suggest that only the neutrons in light nuclei contribute appreciably to π^0-production in proton bombardment. However, differences in the nuclear structure of these nuclei may modify this simple inference.

* This work was performed under the auspices of the AEC.
† Now at the University of Chicago, Chicago, Illinois.
‡ Now at the University of Illinois, Urbana, Illinois.
[1] Steinberger, Panofsky, Steller, Phys. Rev. **78**, 802 (1950).
[2] Bjorklund, Crandall, Moyer, and York, Phys. Rev. **77**, 213 (1950).
[3] W. Crandall and W. Panofsky, private communication; Crandall, Crowe, Moyer, Panofsky, Phillips, Walker (to be published).
[4] Hildebrand, Knable, and Leith, University of California Radiation Laboratory Report No. 1517 (1951), to be published in Rev. Sci. Instr.

Neutral V-Particle Decay and the Negative Proton*

L. I. SCHIFF
Stanford University, Stanford, California
(Received November 28, 1951)

RECENT experimental evidence on neutral V-particles[1] indicates that one variety (denoted here by V^0) decays into a proton and a negative meson (probably pi). The V^0-particles observed thus far are accompanied on the average by several penetrating particles (protons and pi-mesons) and originate in events the total energy of which averages several times 10^{10} ev. Substantially lower energy events seem much less likely to produce V^0-particles, and substantially higher energy events result in penetrating showers of such high density that V^0-particles are hard to identify.

It is reasonable to inquire why it is that a neutral V-particle that decays into a negative proton and a positive meson does not also appear with comparable probability (such a particle will be denoted here by $\bar{V}^0$). V^0 and $\bar{V}^0$ would be expected to have the same rest mass, and indeed to be antiparticles of each other. It can then be argued that they should be produced equally readily, whether they are made independently or as members of antiparticle pairs. The disagreement between this argument and the experimental results to date can be resolved either by supposing that $\bar{V}^0$-particles, and presumably then also negative protons, do not exist, or by supposing that V^0- (or $\bar{V}^0$-) particles can actually be produced singly by some kind of coalescence of a proton (or negative proton) and a negative (or positive) meson. The first supposition would require a drastic revision of the quantum theory of particles as it now exists, and should not be adopted until the evidence for it is unambiguous. The second supposition suffers from the well-known difficulty of reconciling the relatively long lifetime for V^0-decay and the relatively high probability for V^0-production. Apart from this objection, to which we return later, it would account for the experiments in that a $\bar{V}^0$-particle could not then appear unless a negative proton were made as an intermediate step; this would necessitate the simultaneous production of an extra proton or V^0-particle, so that the whole process would require approximately 2 Bev more energy in the center-of-mass system than the appearance of a V^0-particle.

The second supposition can be put in explicit form by making the plausible assumption that there is a conservation law for all particles of nucleonic mass: protons, neutrons, V^0-particles, and other possible particles that decay into protons or neutrons, such as perhaps the charged V-particles. This is a simple extension of the generally accepted conservation law for protons and neutrons only, and may be stated as follows: the difference between the total number of nucleonic particles (protons, neutrons, V^0) and the total number of antinucleonic particles (negative protons, antineutrons, $\bar{V}^0$) is a constant of the motion. This constant is, for example, equal to two if the primary event is the collision of a high energy proton or neutron with a single nuclear proton or neutron. From the point of view of Fermi's statistical treatment of high energy events,[2] an additional restraint is thereby imposed on the system, and a new parameter analogous to the chemical potential in conventional thermodynamics is introduced.

In a two-body encounter, the total number of protons, neutrons, and V^0-particles that emerge should be two unless the energy is so high that antinucleonic particles can be created. As pointed out by Fermi, this will require a primary proton energy greater than about 10^{12} ev, in which case a penetrating shower of high density is produced and antinucleonic particles are difficult or impossible to identify. On the other hand, it is known that several nucleonic particles (including in one instance three neutral V-particles) often emerge from events where the total energy probably does not exceed 10^{11} ev.[3] According to the view adopted here, such events would be regarded as cascade processes that probably take place within a single nucleus.

Sachs[4] has recently suggested that the long V^0-decay lifetime can be understood in terms of the great complexity of the ground and excited states of protons and neutrons (V^0 is supposed to be an excited state of a neutron), which is caused by the presence of many virtual pi-mesons. Such a model can also be used to account qualitatively for the relative ease with which V^0-particles are produced in energetic nuclear encounters, since the high temperature in the small collision volume[2] greatly increases the populations of states which contain large numbers of real and virtual mesons in comparison with the populations of these states for an isolated nucleon or V^0-particle. It seems to us, however, that the essential feature required for an explanation of the apparent lack of reversibility between V^0-formation and decay is independent of the particular model proposed by Sachs, and has to do with the existence of states that connect V^0 and its decay products that are only present with appreciable probability in a region of high temperature.

The writer is greatly indebted to Professors W. B. Fretter, R. B. Leighton, and C. D. Anderson for discussion of their most recent experimental findings.

* Assisted in part by the joint program of the ONR and AEC.
[1] Thompson, Cohn, and Plum, Phys. Rev. **83**, 175 (1951); Leighton, Wanlass, and Alford, Phys. Rev. **83**, 843 (1951); W. B. Fretter, Phys. Rev. **83**, 1053 (1951); private communications from R. B. Leighton and W. B. Fretter; Armenteros, Barker, Butler, and Cachon, Phil. Mag. **42**, 1113 (1951); the writer is indebted to Professor Fretter for showing him prior to publication the manuscript of the paper last cited.
[2] E. Fermi, Prog. Theor. Phys. **5**, 570 (1950).
[3] Private communication from R. B. Leighton.
[4] R. G. Sachs, Phys. Rev. **84**, 305 (1951).

The Saturation Magnetic Moment of Alloys on the Collective Electron Theory*

J. E. GOLDMAN
Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received December 5, 1951)

BOZORTH[1] has called attention to the fact that for certain alloy systems, the Pauling curve for the saturation magnetic moment of the ferromagnetic metals and alloys of the first transition series does not apply. Specifically for the alloys Ni-Mn, Ni-Cr, Ni-V, Co-Mn, and Co-Cr, the addition of small amounts of the nonferromagnetic constituent causes the average moment of the alloy to decrease by anomalously large values instead of the increase predicted by the Pauling curve. A qualitative explanation of this anomaly on the basis of the Heitler-London-Heisenberg approach to the exchange interaction using the Néel criterion[2] has previously been proposed by the author for the case of Ni-Mn[3] and is readily extended to the other four alloys. Néel's condition for ferromagnetism is determined by the difference between the interatomic spacing of the interacting neighbors and the sum of the radii of the magnetic shells. If one uses the radii calculated by Slater,[4] then on the Néel theory, the interaction in the case of all five of the alloy combinations cited is negative, i.e., antiferromagnetic. One might suppose, therefore, on this localized electron picture that the atomic moment of the solute atom is oppositely directed to that of the ferromagnetic matrix. Reasonable agreement in magnitude can then be obtained if one uses an atomic moment consistent with Hund's rule of maximum multiplicity within the atom, allowing for the same number of noncontributing conduction electrons as in the ferromagnetic elements. Zener[5] and his collaborators have given a somewhat analogous interpretation, in which the interaction is again assumed to be an antiferromagnetic one but originating in the probable role of the conduction electrons as originally set forth in Zener's theory.[5] Both these theories have in common that they use a strictly localized electron (Heitler-London-Heisenberg) approximation. The purpose of this note is to propose an alternative explanation in terms of the collective electron model and to suggest an experiment which may permit one to decide between these two approaches to ferromagnetism—at least as applied to the saturation moment of such alloys.

Consideration of the effect of a perturbation in the periodic potential due to a foreign atom on the energy bands of the matrix has led Slater, James,[6] Jones,[7] and others to the conclusion that the band is itself perturbed in such a way that localized in the region near the impurity are allowed levels with energy values, which in the nonperturbed case would normally lie in the forbidden range. For the case of an impurity atom of lower atomic number these states are above the band, and there results a region of depleted charge in the neighborhood of the impurity atom. The argument is readily carried over to the case of a *d*-band in transition elements, with the *d*-electrons considered as itinerant. Thus the *d*-band in cobalt is altered by the presence of a Cr atom in the manner shown schematically in Fig. 1, which corresponds to Slater's picture of an Al impurity in Ge except that the spectrum of discrete states is shown as a continuum. Stated otherwise, the probability of finding an electron of low energy (near the bottom of the band) in the neighborhood of the impurity is small due to the higher potential of the impurity atom with a consequent depletion of charge in that region. Thus the density of states at the bottom of the band is decreased. The redistribution of the electrons among the available states means that in order to minimize the Fermi energy some previously empty states of negative spin (opposite to the macroscopic moment) be occupied. The extent of the decrease in moment is determined by the magnitude of the perturbing potential and should be greater as the atomic number of the impurity is decreased, in agreement with observation.

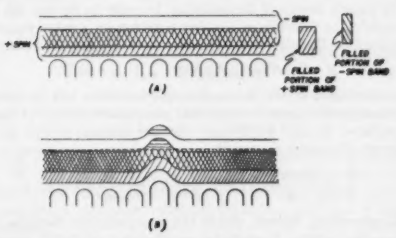


FIG. 1. Schematic representation of the perturbing effect of an alloying element on the energy bands in a metal.

This now suggests an experiment, the result of which will enable one to distinguish between the two approaches previously outlined. If one were to study the susceptibility as a function of temperature of alloys containing enough dissolved nonferromagnetic constituent to reduce the saturation moment to zero, then according to the atomic viewpoint these should be antiferromagnetic, whereas on the collective electron theory one is more likely to find either diamagnetism or a modified Pauli paramagnetism similar, perhaps, to that in copper-nickel alloys with copper concentrations greater than 60 percent in which anomalous susceptibilities have been observed and have been given an analogous interpretation by the author.[8] Such an experiment is currently in progress in this laboratory. The author is indebted to Professor R. Smoluchowski and Dr. N. Rostoker for helpful criticism.

* Supported by the ONR.
[1] R. M. Bozorth, Phys. Rev. **79**, 887 (1950).
[2] L. Néel, Proc. of the Strasbourg Conference on Magnetism Ferromagnetism II, 79 (1940).
[3] J. E. Goldman, J. Appl. Phys. **20**, 1131 (1949).
[4] J. C. Slater, Phys. Rev. **36**, 57 (1930).
[5] W. J. Carr, Jr. (to be published). The author is indebted to Dr. Zener for discussing these results in advance of publication.
[6] J. C. Slater, Phys. Rev. **76**, 1592 (1949); H. M. James, Phys. Rev. **76**, 1602, 1611 (1949).
[7] H. Jones, Bristol Solid State Summer Seminar (September, 1951).
[8] J. E. Goldman, Phys. Rev. **82**, 339 (1951).

The Electronic Specific Heat in Chromium and Magnesium*

S. A. FRIEDBERG, I. ESTERMANN,† AND J. E. GOLDMAN
Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received December 5, 1951)

THE specific heats of several metals, including some in the first transition series, have been measured as a function of temperature between 1 and 4°K. The measurements permit an accurate calculation of the electronic contribution to the specific heat and evaluation of the density of energy states at the top of the Fermi distribution. The cryogenic technique and details of the measurement will be reported elsewhere.[1] The purpose of this note is to call attention to some of the results and their relations to certain fundamental properties of these metals.

The specific heat of metals can generally be represented by a relation of the form $C_v = \gamma T + \beta T^3$. The linear term in temperature represents the electronic contribution, and the coefficient γ gives a direct measure of the density of states at the Fermi surface; the second term is the well-known Debye term that is present in insulators as well as in metals. The two may readily be separated at very low temperatures. This is conveniently done by plotting C_v/T against T^2. The intercept then gives γ directly and the slope gives the Debye characteristic temperature. Our results for Mg and Cr plotted in this manner are given in Fig. 1. The values of γ thus obtained, as well as those for Ti and Zr measured in this laboratory, are compared with measurements on other metals in Table I. The data for V are determined from the critical field data obtained on superconducting vanadium.[2]

The very high values of γ for the transition elements are associated with contributions to the specific heat from the highly dense

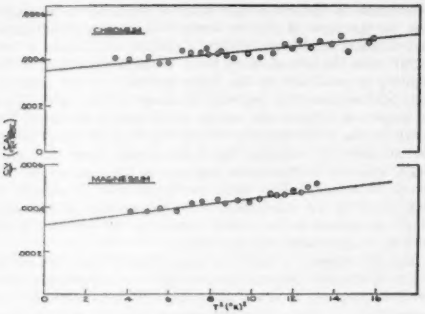


FIG. 1. C_v/T as a function of T^2, for Cr and Mg.

incomplete 3*d*-band. The value in Cr thus appears anomalously low. This is particularly surprising in view of the remarkable similarity between Cr and V, both of which have identical crystals structures with nearly identical spacing and are neighbors in the periodic table. A possible interpretation may be given this phenomenon in terms of the *d*-band shape calculated by Slater and Krutter[3] and by Fletcher and Wohlfarth[4] for the case of Cu, which shows a minimum in the density curve when the band is about half-full. It is not likely that this shape would change significantly as the number of *d*-electrons is decreased. Also, the change in crystal structure is probably not significant, as shown by Manning[5] who finds only minor differences between the Slater-Krutter curve of density of states *vs* energy for Cu and his own for body-centered Fe. In Cr, the Fermi level may be close to this minimum. Another possibility involves the suggestion recently put forward by Zener,[6] that Cr is probably antiferromagnetic and that in a case where the *d*-band is half-full (as in Cr) transitions involving electronic 3*d*-states would require an energy greater than the exchange energy since an electron would have to reverse its spin during the transition. In an antiferromagnet, the band is split in two,[7] and in this case the lower band is full and the upper one empty. This supposition is apparently confirmed by the high temperature specific heat measurements of Armstrong and Grayson-Smith,[8] who observe a rapid increase in C_v above the Dulong-Petit value at temperatures well below the degeneracy temperature, which can be understood if at those temperatures kT is approaching the magnitude of the exchange energy where this anomalous rise begins. On the other hand, the inability of Shull[9] to find evidence for marked antiferromagnetism in Cr, or for a significant magnetic scattering for the chromium atom, tends to support the former interpretation.

The case of magnesium is also of particular interest. From Fig. 1 it is evident that the electronic contribution is very well represented by a linear temperature term. In this material, several investigators have observed a minimum in the curve of resistance *vs* temperature at temperatures between 5° and 15°K. One of the possible interpretations given for this phenomenon has been that there is a small but finite energy gap between the 3*s*- and 3*p*-bands which causes the material to behave as a semiconductor at very low temperatures. In that case, the electronic terms in the specific heat would come from the tail end of the Fermi distribution which is essentially a classical distribution and should therefore depend exponentially on temperature. Our results preclude this interpretation. This is confirmed by estimates of the effective mass of the conduction electrons based on our measured density of states and the magnetic susceptibility data of McGuire,[10] which gives a result considerably greater than the free electron mass. The authors are indebted to Professor J. C. Slater for pointing out Shull's results on chromium as well as for discussions of the band theory interpretation. We are also indebted to Dr. Clarence Zener for discussions of his theory of the properties of transition elements.

TABLE I. Values of γ, the coefficient of the term linear in T in the specific heat; and $n(\epsilon)$, the density of states at the Fermi surface.

Metal	$\gamma \times 10^4$	$n(\epsilon)$ levels/atom/ev
Copper	1.80	0.16
Magnesium	3.25	0.29
Titanium	8.3	0.74
Vanadium	14	
Zirconium	6.92	0.62
Chromium	3.74	0.33

* Supported by the ONR and the Research Corporation.
† Presently on leave at the Office of Naval Research, Washington, D. C.
‡ S. Friedberg, thesis, Carnegie Institute of Technology (1951); I. Estermann and S. Friedberg (to be published).
§ We are indebted to Dr. A. Wexler for calling these results to our attention.
[3] H. M. Krutter, Phys. Rev. **48**, 664 (1935); J. C. Slater, Phys. Rev. **49**, 537 (1936).
[4] G. C. Fletcher and E. P. Wohlfarth, Phil. Mag. **42**, 106 (1951).
[5] M. F. Manning, Phys. Rev. **63**, 190 (1943); J. B. Greene and M. F. Manning, Phys. Rev. **63**, 203 (1943).
[6] C. Zener, Phys. Rev. **81**, 440 (1951).
[7] J. C. Slater, Phys. Rev. **82**, 538 (1951).
[8] L. D. Armstrong and H. Grayson-Smith, Can. J. Research **A28**, 51 (1950).
[9] C. G. Shull, private communication.
[10] T. McGuire, private communication. We are indebted to Dr. McGuire for making available to us the results in advance of publication.

A Relativistic Modification of Bose and Fermi Statistics

A. E. SCHEIDEGGER AND C. D. MCKAY
Queen's University, Kingston, Ontario, Canada
(Received November 23, 1951)

ONE of the most striking results in our previous investigations[1] on the quantum statistics of fields was that the thermodynamics of the relativistic Schrödinger-Gordon field does not yield the thermodynamics of the nonrelativistic field if c (the velocity of light) approaches infinity. One can show that this is the result of the fact that the rest energy of the particles is included in the Hamiltonian of the relativistic field, whereas in the nonrelativistic case it is not. This is most easily seen if the average energy per particle is calculated. A straightforward computation by means of the results of our earlier publication yields for $kT \ll mc^2$

$$E/N \simeq mc^2 + \tfrac{3}{2}kT. \tag{1}$$

It seems therefore desirable to develop a subtraction formalism which effects that the thermodynamic functions of the relativistic field refer only to that part of the energy whose classical analog is the kinetic energy of the particles. In order to do this one has to replace the Schrödinger-Gordon equation by a wave equation such that the Hamiltonian of the quantized field may be written as

$$H = \Sigma_K N_K \hbar c[(\mu^2 + K^2)^{\frac{1}{2}} - \mu]; \quad \mu = mc/\hbar. \tag{2}$$

The calculation of the thermodynamic functions will be exactly as in our earlier paper,[1] except that everywhere $(\mu^2+K^2)^{\frac{1}{2}}$ has to be replaced by $[(\mu^2+K^2)^{\frac{1}{2}}-\mu]$. For the internal energy E, for instance, this yields the result

$$\frac{E}{V} = \frac{3}{2}\frac{(2\pi m)^{3/2}}{h^3}(kT)^{5/2}\Sigma\frac{1}{n^{3/2}}\left\{1+\frac{75}{16}\frac{kT}{mc^2}+\cdots\right\}. \tag{3}$$

This expression, indeed, yields the nonrelativistic formula if c approaches infinity. It might be regarded as a relativistic generalization of Bose statistics; the correct relativistic energy resulting

from the velocity of the particles is taken thermodynamically into account, but not the energy contained in the rest mass of the particles. Therefore Eq. (3) would correspond to the statistics of an ensemble of particles which are lifted out of an unobservable reservoir without thermal energy being used for their creation, but their motion being described relativistically.

The subtraction of the rest energy of the particles representing the field may seem somewhat arbitrary. It is interesting to notice that a generalization of ordinary statistical mechanics effects the same result. In order to show this we have to go back to the very principles of statistical thermodynamics. The fundamental problem in that theory is to determine the distribution of an ensemble of a large number of identical systems over the possible states in which this ensemble can find itself, given that the energy of the ensemble is a constant. However, it has long been known that the distinguished role of the energy is simply that it is a constant of the motion. It is possible to generalize the formalism so as to refer to other constants of the motion as well. In this instance the partition function can be written

$$Z = \text{trace} \exp[-(1/k)\Sigma\alpha_i \mathbf{A}_i], \tag{4}$$

where the $\mathbf{A}_i$ are operators corresponding to arbitrary classical constants of the motion, and where the α_i are their respective inverse "temperatures."

With these notions it is possible to extend the thermodynamics of the Schrödinger-Gordon field. Let us take for $\mathbf{A}_1$ the ordinary Hamiltonian, for $\mathbf{A}_2$ the operator $\mathbf{N}$ which represents the number of particles "present" in the field (and corresponds thus to a classical constant of the motion). Then the partition function becomes

$$Z = \text{trace} \exp(-\vartheta\mathbf{H} - k^{-1}\alpha\mathbf{N}). \tag{5}$$

It is obvious that a suitable choice of the "number-temperature" $T_N = 1/\alpha$ will reduce this partition function to that one which one obtains using the Hamiltonian given in (2). From a statistical standpoint, there is no *a priori* reason why the conventional partition function should be preferred over Eq. (5). If T_N in (5) is allowed to vary, different thermodynamic functions (with respect to the ordinary temperature $T = 1/k\vartheta$) are found for every value of it. Again, there is no *a priori* reason for one set of such thermodynamic functions being preferred before another. The choice has to be made on physical grounds. Thus it seems that the set of formulas which were obtained in our earlier paper and the thermodynamic functions discussed here are both perfectly correct thermodynamic descriptions of the Schrödinger-Gordon field. Which set of formulas has to be used in a particular case depends solely upon the nature of the hypothetical interactions which are supposed to produce the canonical distribution of the fields.

In conclusion, one should perhaps remark that very similar facts apply when the Dirac field and its nonrelativistic form are studied. In this case the thermodynamic functions lead, after the subtraction of the rest mass of the particles, to a relativistic modification of Fermi statistics.

[1] A. E. Scheidegger and C. D. McKay, Phys. Rev. 83, 125 (1951).

The Half-Life of Positrons in Condensed Materials*

S. DeBenedetti and H. J. Richings
Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received November 28, 1951)

THE difference in the half-lives of positrons in several pairs of condensed materials was experimentally studied with the method of delayed coincidences between the nuclear γ-rays of Na^{22} and the annihilation radiation from the positrons of the same source.[1]

Two stilbene counters were pointed at the source, at about 90° from each other. The biases were so adjusted that a coincidence

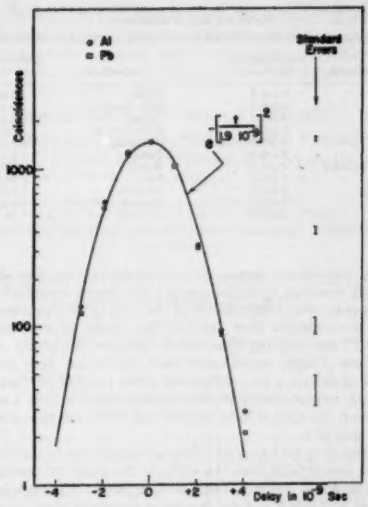


FIG. 1. Comparison of positron half-lives in Al and Pb.

was recorded only if a 1.3-Mev γ was detected by one of the counters and an annihilation γ by the other.

The Na^{22} source S was deposited on thin Al and sandwiched between the substances A and B for which the half-life was compared. With a source sandwich arranged in the order $BASAB$ the positrons stopped in A, while if the sandwich was arranged as $ABSBA$, the positrons stopped in B. However, the absorption of the 1.3-Mev γ-rays of Na^{22} and of the annihilation radiation was essentially the same in the two cases; in this way directly com-

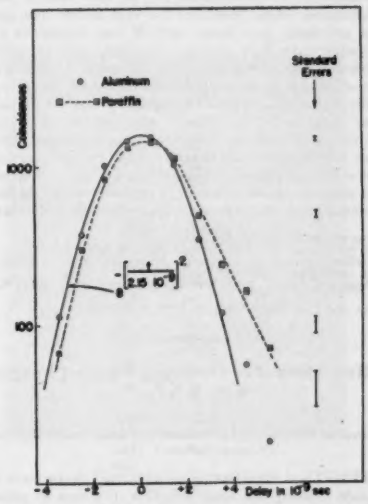


FIG. 2. Comparison of positron half-lives in Al and paraffin.

TABLE I. Difference in the positron half-life in the substance indicated and in aluminum.

Substance	Difference in 10^{-10} sec	Substance	Difference in 10^{-10} sec
Li	0 ± 0.5	Al_2O_3	0 ± 0.5
Na	0 ± 0.5	SiO_2 (quartz)	2.5 ± 1
K	0 ± 0.5	Teflon	1.2 ± 0.5
Ag	0 ± 0.5	Polystyrene	2.5 ± 1
Pb	0 ± 0.5	Plexiglass	2.5 ± 1
Cu	0 ± 0.5	Paraffin	2.5 ± 1
C (graphite)	0 ± 0.5	Water	1.5 ± 0.7
S	0 ± 0.5	H_3O_3	0 ± 0.5
NaCl	0 ± 0.5	Benzene	2.6 ± 1
AgCl	0 ± 0.5	Alcohol (methyl)	2.3 ± 1
PbO	0 ± 0.5		

parable coincidence curves were obtained and spurious delays (possibly resulting from changes in pulse height connected with difference in energy degradation of the γ-rays) were avoided.

The coincidences were selected by means of a circuit of $\approx 2\times10^{-9}$ sec resolving time, which had been previously tested with time-of-flight experiments[2] and which had been proved capable of detecting time differences of the order of 10^{-10} sec.

Typical coincidence curves obtained are shown in Figs. 1 and 2. Aluminum was used in most cases as one of the substances in the source sandwich.

By assuming that positron lives are exponentially distributed in time, one obtains from the shifts in the empirical curves the results shown in Table I. There is no apparent relation between the data of Table I and the density or atomic number, as might be expected if the delays were a result of slowing-down time. Thus, the results have probably to be interpreted in terms of differences in annihilation half-lives.

It is somewhat surprising that all metals tested exhibit the same half-life, for this is in contradiction with the simple view that positrons annihilate at low energy in free collision with the valence electrons.[3] From this viewpoint measurable differences should occur at least in the alkali series owing to the wide variation in atomic density. It is tempting to think that the positron and the conduction electrons form some kind of fast annihilating state, thus making the lifetime independent of the metal used.

In the case of insulators, one might be tempted to look for some connection between positron half-life and chemical properties. Substances whose molecules are very stable, like quartz, paraffin, and water, have longer half-life than substances which more readily enter into chemical reactions, such as sulfur or hydrogen peroxide. It is possible that this might be evidence for the formation of positron compounds in condensed materials.

Finally, one should consider the possibility that the time shifts observed were caused by different ratios of singlet to triplet annihilation in different materials. A greater percentage of three-quantum annihilation[4] could be accompanied by a longer positron life, but would also produce apparent time shifts as a result of the delayed response of our instrument to lower energy radiation. This point will be investigated experimentally in the near future.

* Supported in part by the AEC.

[1] J. W. Shearer and M. Deutsch, Phys. Rev. **76**, 462 (1949).

[2] S. DeBenedetti and H. J. Richings, Phys. Rev. **82**, 771 (1951); and Rev. Sci. Instr. (to be published).

[3] DeBenedetti, Cowan, Konneker, and Primakoff, Phys. Rev. **77**, 205 (1950).

[4] A. Ore and J. L. Powell, Phys. Rev. **75**, 1696 (1949).

A New Method of Obtaining Focused Images with X-Rays

G. N. RAMACHANDRAN
Department of Physics, Indian Institute of Science, Bangalore, India
(Received October 1, 1951)

ATTEMPTS at constructing an x-ray microscope have been made utilizing the total reflection of x-rays at glancing angles less than about 30 minutes of arc.[1-3] Since the images are highly astigmatic, a combination of two concave mirrors with their planes of reflection at right angles is necessary. Magnifications of 15 to 30 have been reported with an object of dimensions of approximately 0.4 mm square.

It occurred to the author that under certain conditions, focused images could be obtained using crystal reflections. The use of Bragg reflection in various types of cameras for focusing a particular wavelength to a linear focus is familiar to the x-ray crystallographer. In the method to be described below, the wavelength of the x-rays may vary, but rays diverging from a particular point are again brought to a focus at another point.

Crystal reflection has the property that the angle of incidence is equal to the angle of reflection for a particular set of lattice planes, but for a given value of this angle θ only a narrow range of wavelengths on either side of a wavelength λ satisfying the Bragg relation $\lambda = 2d\sin\theta$ is reflected. Thus those rays which are reflected, are reflected according to the same law as ordinary light, incident at the same angle on a plane mirror placed in the same orientation as the lattice plane of the crystal under consideration.

Now in the optical case, a concave mirror would give a focused image (either reduced or magnified) of an object placed in front of it at a suitable distance. If the mirror were now replaced by a thin cleavage plate of the crystal, curved to the same surface as the mirror, and if white x-rays were allowed to illuminate the object, then according to the argument of the preceding paragraph one should get a focused image with x-rays of the same size and at the same place as that produced in the optical case. It is important to realize that of all the rays which originate from a point on the object, those which are reflected must converge to the corresponding point of the image, if the mirror has no aberrations. The others are just transmitted. Thus the only background would be that resulting from x-rays incoherently scattered by the mirror.

This very simple possibility does not appear to have been pointed out before. Experiments have been undertaken by the author in collaboration with Mr. Y. T. Thathachari in this laboratory to test the preceding ideas, and it is found that they are workable. Although details of the experiment will be published elsewhere, the possible methods of constructing such an x-ray focusing arrangement may be mentioned here. One method would be to fix a cleavage flake of mica on a flange attached to a circular tube and to partially evacuate the inside of the tube, thereby obtaining a concave surface. Another method would be to bend a circular plate of quartz, cut perpendicular to the c-axis, by means of two concentric rings, where the region inside the smaller ring should have nearly constant curvature.[4] A third method would be to deform a plastic crystal like rocksalt to take the shape of a truly spherical surface. In all these cases, one would expect the lattice planes to take up the requisite curvature (at least on a macroscopic scale).

The geometry of the arrangement would be as follows. The object, kept inside an aperture on a lead screen, is placed close to the window of an x-ray tube and the diverging beam of x-rays is allowed to fall on the "mirror" situated at a suitable distance from it. The mirror is slightly tilted so as to throw the reflected beam on one side of the object, and a photographic plate is kept at the correct distance from the mirror, judged optically. The x-ray image also would occur at the same place and could thus be recorded.

This technique has interesting possibilities in connection with constructing an x-ray microscope. Although the image would be astigmatic in the aforementioned arrangement, it would be possible to get an axial arrangement by combining two mirrors (with central holes) similar to the arrangement in an optical reflection microscope. But, unlike the optical case, the wavelength of a ray reflected by the first mirror is dependent on its angle of incidence, and very special conditions have to be satisfied by the second mirror if each ray is again to be incident at the Bragg angle on it. These conditions have been worked out, and

the results of the theory will be published elsewhere. It is found that the accuracy in setting of the two mirrors is of the same order as that required in a double crystal spectrometer.

I am grateful to Mr. S. Ramaseshan and Mr. G. Suryan for the discussions I had with them on this subject.

[1] W. Ehrenberg, Nature **160**, 330 (1947).
[2] P. Kirkpatrick and A. V. Baez, J. Opt. Soc. Am. **38**, 766 (1948).
[3] C. M. Lucht and D. Harker, Rev. Sci. Instr. **22**, 392 (1951).
[4] S. Timoshenko, *Theory of Plates and Shells* (McGraw-Hill Book Company, Inc., New York, 1940).

Antiproton Annihilation*

ISRAEL REFF
Physics Department, Indiana University, Bloomington, Indiana
(Received November 16, 1951)

THE cross section for meson production by proton-antiproton annihilation was calculated by the usual second-order perturbation procedure. The results for neutral s.-s. (scalar theory with scalar coupling), charged s.-s., neutral p.s.-p.s., and charged p.s.-p.s. are in agreement with those of Ashkin, Auerbach, and Marshak.[1] Results for the other cases are given in Tables I and II for the center-of-mass system in units of $\hbar=1$, $c=1$, for the limit of $p\to 0$ (p = nucleon momentum) and $p\to\infty$ (extreme relativistic limit, nucleon energy greater than five times nucleon rest mass). K = nucleon mass; μ = meson mass; θ = angle between proton and positive meson for charged cases, angle between proton and neutral meson for neutral cases

$$A=(2p^2+2K^2-\mu^2)^2-4p^2(p^2+K^2-\mu^2)\cos^2\theta.$$

This problem was suggested by discussions with Professor J. G. Retallack. The author thanks Professor J. V. Lepore for many helpful discussions on this calculation.

TABLE I. Total and differential meson production cross sections for the limit $p\to 0$.

Interaction	Total cross section $p\to 0$	Differential cross section $p\to 0$
neutral p.s.-p.v.	$\frac{4}{3\pi}\frac{g^4K}{\mu^4}p$	$\frac{g^4K}{2\pi^2\mu^4}p\sin^2\theta$
charged p.s.-p.v.	$\frac{g^4}{8\pi}\left(\frac{K}{\mu}\right)^4\frac{1}{Kp}$	isotropic
neutral v.-v.	$\frac{g^4}{144\pi}\frac{1}{Kp}$	isotropic
charged v.-v.	$\frac{g^4}{72\pi}\left(1+\frac{9}{2}\frac{\mu^2}{K^2}\right)\left(\frac{K}{\mu}\right)^4\frac{1}{Kp}$	isotropic
neutral p.v.-p.v.	$\frac{g^4}{144\pi}\frac{1}{Kp}$	isotropic
charged p.v.-p.v.	$\frac{g^4}{72\pi}\left(1-\frac{3}{2}\frac{\mu^2}{K^2}\right)\left(\frac{K}{\mu}\right)^4\frac{1}{Kp}$	isotropic

TABLE II. Total and differential meson production cross sections for the limit $p\to\infty$.

Interaction	Total cross section $p\to\infty$ Extreme rel. limit	Differential cross section $p\to\infty$ Extreme rel. limit
neutral p.s.-p.v.	$\frac{g^4K^2}{\pi\mu^4}\left(2+\frac{K^2}{p^2}\ln\frac{4p^2}{K^2}\right)$	$\frac{g^4K^2}{2\pi^2\mu^4}\left(1+\frac{4K^2p^2}{A}\right)$
charged p.s.-p.v.	$\frac{g^4}{12\pi}\frac{p^2}{\mu^4}$	$\frac{g^4}{32\pi^2}\frac{p^2}{\mu^4}\sin^2\theta$
neutral v.-v.	$\frac{g^4}{144\pi p^2}\left(\ln\frac{4p^2}{K^2}-1\right)$	$\frac{g^4}{64\pi^2}\left(\frac{8p^2}{9A}-\frac{1}{9p^2}\right)$
charged v.-v.	$\frac{g^4}{108\pi}\frac{p^2}{\mu^4}$	$\frac{g^4}{288\pi^2}\frac{p^2}{\mu^4}\sin^2\theta$
neutral p.v.-p.v.	$\frac{g^4K^2}{36\pi\mu^4}\left(2-\frac{K^2}{p^2}\ln\frac{4p^2}{K^2}\right)$	$\frac{g^4K^2}{72\pi^2\mu^4}\left(1-\frac{4K^2p^2}{A}\right)$
charged p.v.-p.v.	$\frac{g^4}{108\pi}\frac{p^2}{\mu^4}$	$\frac{g^4}{288\pi^2}\frac{p^2}{\mu^4}\sin^2\theta$

* Supported by the joint program of the ONR and AEC.
[1] Ashkin, Auerbach, and Marshak, Phys. Rev. **79**, 266 (1950).

Anisotropy in Paramagnetic Resonance Absorption of Picryl-*n*-Amino Carbazyl*

VICTOR W. COHEN AND C. KIKUCHI†
Brookhaven National Laboratory, Upton, New York
AND
JOHN TURKEVICH
Princeton University, Princeton, New Jersey
(Received December 5, 1951)

WE have examined the paramagnetic resonance of a new free organic radical, picryl-*n*-amino carbazyl. A report on its synthesis will appear in print shortly. From arguments based upon its molecular structure, it was expected that its absorption line would be considerably narrower than that of α,α diphenyl-β-picryl hydrazyl as a result of stronger exchange effect.

The observations were made using a K-band cylindrical cavity operated in the TE_{01} mode with the cavity axis normal to the applied magnetic field. The specimen was cemented to the end of a quartz rod which protruded along the axis of the cavity to its midpoint. The specimen could then be rotated easily about the cavity axis. The absorption lines were observed by sweeping the magnetic field while the frequency was kept constant. The sensitivity was such that a sample of 5×10^{-4} cm³ carbazyl could produce a good trace on the oscilloscope.

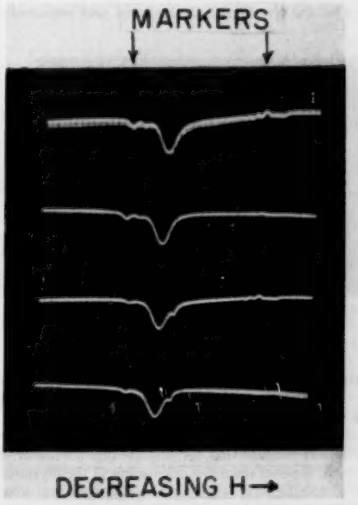


FIG. 1. Carbazyl and hydrazyl samples, showing shift of carbazyl peak with rotation.

For a sample which obviously consisted of several crystals the line width was approximately 7 gauss. This line showed a marked structure, in which the components shifted rapidly as the sample was rotated. Going to a much smaller sample, the spectrum was found to consist of a few clearly resolved lines, each approximately 0.5 gauss wide. The position of the line appears to depend upon the orientation of the crystal varying over a range of 7 gauss upon rotating the sample through 90°.

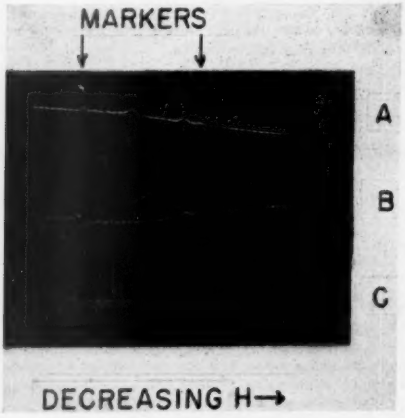


FIG. 2. Absorption spectrum of carbazyl showing two lines, one of which is independent and the other dependent upon orientation. (a) Two peaks slightly separated; (b) two peaks in coincidence; (c) rotation 90° from (b).

The g value of the carbazyl was measured relative to that of the hydrazyl. For this purpose a small polycrystalline specimen of the hydrazyl was also mounted on the quartz rod. As the quartz rod was rotated (see Fig. 1), the hydrazyl peak remained stationary, while the carbazyl peak varied from 5 gauss above to 2 gauss below the hydrazyl peak, corresponding to the effective g-values 2.0024 and 2.0041 respectively, assuming the value of 2.0036 for hydrazyl.[1]

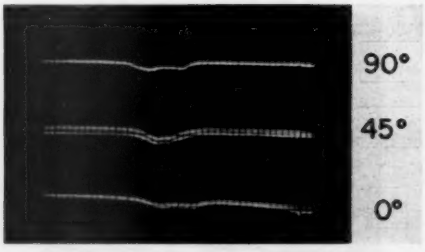


FIG. 3. Hydrazyl lines observed for three orientations. (a) Arbitrary reference; (b) 45°; (c) 90°.

Our results indicate that the angular dependence of the resonance peak depends upon the orientation of the crystalline axes. In one sample (see Fig. 2), the absorption spectrum consisted of two lines, one of which showed no directional effect and the other depended strongly on the orientation of the sample. These results seem to indicate that the observed angular dependence is caused by molecular diamagnetism.

A similar effect was observed for single crystals of hydrazyl[3] (see Fig. 3). In this case the peak varied over a total range of 5 gauss, or about $\frac{1}{7}$ that of carbazyl.

* Work done under contract with the AEC.
† On leave from Michigan State College, East Lansing, Michigan.
[1] Holden, Kittel, Merritt, and Yager, Phys. Rev. 77, 147 (1950); C. H. Townes and J. Turkevich, Phys. Rev. 77, 148 (1950).
[3] The sample used in this observation was furnished through the courtesy of Dr. Garstens and Dr. Liebson of the Naval Research Laboratory.

Sublimation and Condensation of Carbon Dioxide, Nitrogen, and Oxygen in an Expansion Chamber*

B. M. CWILONG AND M. H. EDWARDS
Department of Physics, University of British Columbia, Vancouver, Canada

(Received September 10, 1951)

THE question of what actually happens in supersaturated water vapor below the freezing point has aroused considerable interest during recent years. A very sensitive experimental method was developed for detecting the presence of the first ice particles in a droplet cloud.[1] (A cloud is formed over a pool of supercooled water, in an expansion chamber precooled below the freezing point. The first ice particle present inoculates the pool and causes it to freeze, the freezing of the pool thus serves as an indicator.) This method allowed a threshold temperature for the appearance of ice (−41.2°C) to be established.

The question then arose as to whether these ice crystals are originated by sublimation or by freezing of minute water droplets. As condensation of water vapor requires considerable supersaturation, a crucial experiment was sought to separate sublimation from condensation, i.e., to observe precipitation in the sublimation temperature region produced by expansion ratios below the threshold values for condensation. The supercooled water pool could not be used here as a detector because the initial supercooling of a chamber has to be in the neighborhood of −40°C, that is, below the temperature region in which supercooled water in bulk can be maintained easily. Therefore observations had to be visual, a rather difficult task, as the showers at best would be very faint in view of the fact that the amount of water available for sublimation in these conditions is small (saturated water vapor pressure at −40°C is 0.0966 mm Hg).

In 1944 Cwilong[1] observed these ice showers with expansion ratios down to 1.07, considerably less than the threshold expansion ratio (1.29 at −40°C) required to produce condensation on negative ions, and therefore concluded that the sublimation is a primary process. In 1948 d'Albe,[2] using a microscope, failed to see any ice showers using expansion ratios up to 1.1. In 1949 Brewer and Palmer[3] reported that in an improved apparatus (details of the work are not published), which eliminates the possibility of cloud contamination by ice splinters from the walls, they did not see any sublimation at all, but just the continuation of C. T. R. Wilson's negative ion threshold for condensation. In 1950 Salmon and Cwilong[4] repeated the experiments once more, also in an apparatus in which contamination of cloud from the walls is impossible, and in a series of over 300 observations detected genuine sublimation at expansion ratios down to 1.05.

Assuming that sublimation does exist, experimental conditions were sought in which the process could be isolated from condensation and easily demonstrated. Carbon dioxide was thought a suitable cloud-producing substance. High vapor pressure over the solid promised abundant precipitations and a very high vapor pressure of the triple point (3885 cm Hg) excluded the existence of droplets in any cloud which might be formed. Expansions were performed in a chamber precooled to the region of −100°C and containing saturated vapor of CO_2 in a helium atmosphere.

It was found that the expansion ratio 1.2 is a sharp threshold value. Any expansion below this value produces supersaturated vapor of CO_2 without precipitation. Expansions above 1.2 produce abundant showers of CO_2 crystals. The threshold value increases slightly with lowering of the initial temperature of the chamber. There is an appreciable increase in the number of crystals formed with an increase of the expansion ratio up to 1.4. Above 1.4 the density of showers appears constant. All crystals grow sufficiently to have an appreciable rate of falling. The chamber is clear of all crystals within a maximum of 4 seconds after any expansion. Irradiation of the chamber with x-rays has no effect on the sublimation threshold, nor does it appreciably affect the number of particles formed.

The experiments provide straightforward evidence for the existence of a genuine sublimation process and also for the existence of a definite sublimation threshold. They also indicate that this sublimation threshold is defined primarily in terms of an expansion ratio and not of temperature.

With the same apparatus nitrogen and oxygen were investigated. Expansions were performed in a chamber precooled to the region of the nitrogen triple point and containing saturated vapor of the investigated gas in a helium atmosphere.

Both vapors show very low condensation thresholds. With oxygen even the smallest expansions (in a helium atmosphere free from condensation nuclei) produce dense fog. The threshold expansion ratio for condensation of nitrogen is 1.005. Dense fog produced at this threshold value indicates that the number of active condensation nuclei is of greater order of magnitude than the number of ions in the chamber. The fog has a consistency similar to C. T. R. Wilson's fogs obtained with water vapor, with expansion ratios above his cloud threshold (1.38).

Below their triple points both vapors form homogeneous, stable supercooled droplet clouds. The first crystals of solid nitrogen in nitrogen droplet clouds were observed at an expansion ratio of 5. In oxygen even higher expansion ratios, up to 10, failed to produce crystals.

* This work was carried out under a grant from the Defense Research Board of Canada.
[1] B. M. Cwilong, Nature **155**, 361 (1945); Proc. Roy. Soc. (London) **A190**, 137 (1947).
[2] E. M. Fournier d'Albe, Quart. J. Roy. Met. Soc. **75**, 1 (1949).
[3] A. W. Brewer and H. P. Palmer, Nature **161**, 312 (1949).
[4] C. Salmon and B. M. Cwilong, New Zealand Sci. Rev. **8**, 78 (1950).

Photocontrolled Low Frequency Dielectric Dispersion*

J. Ross Macdonald
Armour Research Foundation, Illinois Institute of Technology, Chicago, Illinois
(Received December 5, 1951)

PHOTOCONTROLLED dielectric dispersion has been observed in *F*-centered alkali halide crystals at room temperature. Most of the measurements were carried out on KBr single crystals which had been initially *U*-centered and then colored photochemically with Co^{60}-radiation, but the effect has also been verified with additively colored KBr and KI single crystals.

To allow observation of photoeffects, metallic electrodes were either painted or evaporated on opposite crystal faces and incident light directed through the crystal between the electrodes or through one (half-silvered) electrode into the crystal. No significant differences were observed in the behavior of these two types of specimens. Measurements of crystal admittance with or without illumination were carried out over the range from 25 to 10,000 cps and enabled the complex dielectric constant of the crystals to be computed.

In the dark, it was found that crystal dark-currents were purely reactive within the precision of the measurements. The dark dielectric constant of the KBr crystals was found to be about 5, independent of *F*-center concentration and frequency over the ranges measured. These are expected results; the crystals are good insulators since their ionic conductivity is negligible at room temperature.

On strongly illuminating a crystal with light in its *F* absorption band, it was found that the total crystal current at low frequencies increased by a factor of from 5 to 50 over the dark current and was still largely reactive. This photocurrent depended strongly on *F*-center concentration and light intensity. On determining the real part of the dielectric constant ϵ' and the loss factor ϵ'', it was found that these quantities exhibited dielectric dispersion behavior, with ϵ' a maximum at zero frequency and ϵ'' a maximum in the low audiofrequencies. No appreciable frequency-independent photoconductivity was observed. Typical results found for a KBr crystal with 1.4×10^{18} *F*-centers per cm^3 were a maximum value of ϵ' at zero of 500 and a maximum of ϵ'' of 142 occurring at 85 cps. Applying a theory developed by Fuoss and Kirkwood[1] for large molecules with permanent dipoles having a distribution of relaxation times, it has been found possible to fit experimental ϵ'' and ϵ' curves quite well, whereas equations based on a single relation time will not fit the data at all.

Preliminary measurements of photocurrent as a function of *F*-center concentration and light intensity at a single frequency showed that the photocurrent varied approximately with the square root of these quantities. Further measurements of ϵ' and ϵ'' as functions of frequency have shown, however, that the preceding results were caused by a shift in the frequency at which the maximum of the ϵ'' curve occurred toward higher frequencies with increasing *F*-center concentration or light intensity, without an appreciable change in the maximum values of ϵ' and ϵ'' themselves. These results suggest that the effect may be largely the result of Maxwell-Wagner type losses associated with the photoconductivity of a crystal under illumination. No polarization effects have been observed: the photocurrent obeys Ohm's law up to the maximum field strength applied, 500 volts/cm, and the insertion of a 300-volt dc battery in series with the applied ac voltage has no effect on the results. This work will be more fully discussed in a later paper.

* Research carried out in part under Signal Corps contract.
[1] R. M. Fuoss and J. G. Kirkwood, J. Am. Chem. Soc. **63**, 385 (1941).

$L(L+1)$ Correction in the Spectra of the Iron Group

Giulio Racah
Hebrew University, Jerusalem, Israel
(Received November 12, 1951)

IT was recently pointed out by Trees[1] that the agreement between the theoretical formulas and the experimental data in the d^5s configuration of Mn II and Fe III is greatly improved by the addition of a correction term of the form $\alpha L(L+1)$. We verified that this correction is also very important in many other spectra of the iron group, and that the values of α obtained by least squares from different spectra are very consistent.

It is the purpose of this letter to show that this effect may be explained by the sole mathematical assumption that the deviations from the theoretical formulas in the *n*-electron configurations are the sum of the deviations of the interaction of each couple from the corresponding formulas. It is, however, not easy to justify this assumption, as the linearity property, which holds for every first-order effect, is not expected to hold for the deviations from Slater's formulas, which are generally considered to be the result of second-order effects.

It is known that the spectra of the configurations d^2 do not fit Slater's formulas very well;[2] but if we introduce two correction terms $\alpha\varphi_1(L)$ and $\beta\varphi_2(L)$, where φ_1 and φ_2 are any functions of L, it will be always possible to represent exactly the five terms of d^2, as we have now five free parameters.

Without any loss of generality we may therefore represent the deviations from Slater's formulas in d^2 by

$$2\alpha(\mathbf{l}_1\cdot\mathbf{l}_2)+\beta q_{12},$$

where q is the operator, already defined elsewhere,[3] which vanishes for $L\neq 0$ and equals 5 for the S-state. Then, according to our assumption, the deviations in the configuration d^n will be represented by

$$\sum_{i<k}[2\alpha(\mathbf{l}_i\cdot\mathbf{l}_k)+\beta q_{ik}]=\alpha[L(L+1)-6n]+\beta Q.$$

As Q is different from zero only for the higher terms of every configuration and vanishes or has very small expectation values for almost all the terms which are experimentally known, the second correction has no practical importance, and the only important one is the $L(L+1)$ correction.

It may also be pointed out that this correction explains the constancy of the ratio $(^1S-{}^1D)/(^1D-{}^3P)$ in the configurations $2p^2$ and $2p^4$ of the first period, and predicts that the ratio $(^3P-{}^3D)/(^3D-{}^5S)$ in the configuration $2p^3$ should be 4/9 of the former ratio, in excellent agreement with the experimental data.[4]

[1] R. E. Trees, Phys. Rev. **83**, 756 (1951).
[2] E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra* (Cambridge University Press, London, 1935), p. 203.
[3] G. Racah, Phys. Rev. **62**, 438 (1942), Eq. (90), and **63**, 367 (1943), Eq. (35).
[4] See reference 2, p. 198.

The $L(L+1)$ Correction to the Slater Formulas for the Energy Levels

R. E. Trees*
University of Pennsylvania, Philadelphia, Pennsylvania
(Received November 23, 1951)

In recent papers[1,2] it has been pointed out that a correction of form $\alpha L(L+1)$ produces greatly improved agreement between theoretical and experimental term values in the configuration d^5s of Mn II and Fe III. This correction represents a polarization energy, and rigorous treatment of second-order effects is difficult even for the highly simplified wave functions of electrons in a solid.[3] No satisfactory simple theory to explain this effect has as yet been found, but the following comments may be helpful.

The fact that the errors in all configurations of type $3d^n$ decrease as L increases has already been noted.[2] A variation independent of spin and proportional to $L(L+1)$ was, however, felt to be quantitatively justifiable only for the d^5 configuration. In the d^6 configuration of Fe III a correction proportional to $L(L+1)$ reduces the mean deviation only from ± 876 cm^{-1} [4] to ± 439 cm^{-1}.[5] This is a relatively small improvement compared to the reduction from ± 852 cm^{-1} to ± 105 cm^{-1} found in the d^5s configuration of Fe III. The value of α in the d^6 configuration, namely 67.4, is consistent with the value 80.7 in the d^5s configuration. This, together with the similarity of the trend in the errors for all $3d^n$ configurations, indicates that the corrections are at least approximately related according to the theory just advanced by Racah[6] in which the polarization energy is treated linearly. However, as his theory leads essentially to a correction of form $L(L+1)$ for all configurations, it will not realize the full possibilities for accurate prediction of levels to ± 100–200 cm^{-1} in the $3d^6$ configuration of Fe III.

A basis for the favored position of the d^5 configuration relative to the d^6 configuration exists in the fact that in the former case most of the term values are rational functions of the Slater integrals (i.e., the eigenfunctions are characterized by a single seniority number,[7] whereas most of the d^6 eigenfunctions are linear combinations of states with two seniority numbers). The one term with an irrational theoretical value in the $3d^54s$ configuration of Fe III for which an unperturbed experimental value is available is the $3d^5(a^3D)4s$, c^3D. For this term the error was $+442$ cm^{-1} while for all other levels the absolute value of the errors was less than 200 cm^{-1}. It seems likely from this that the correction of form $\alpha L(L+1)$ applies only to terms that are rational functions of the Slater parameters.

This conclusion can be checked very nicely in the $3d^6$ configuration of Fe III against the 6 rational levels for which experimental values are available. A correction proportional to $L(L+1)$ need not be applied to the usual Slater formulas, as such a correction can be absorbed in the values of the parameters A, B, C.[8] These 6 levels can be fitted with a mean deviation of ± 150 cm^{-1}, and no deviation exceeds 200 cm^{-1} in absolute value.[9] This is to be compared with the deviation of ± 876 cm^{-1} in fitting all levels.

The correction to be applied when the term value is not a rational function of the Slater parameters, so that agreement of equal accuracy will be obtained, has not yet been found. However, if the eigenfunction is

$$C_1\psi(d^nv_1SL)+C_2\psi(d^nv_2SL),$$

a correction of form

$$(1+2C_1C_2)\alpha L(L+1)$$

will produce close agreement in the $3d^5(a^3D)4s$, c^3D term discussed previously. A check of the validity of this correction in the d^6 configuration has not yet been made. In any event, it does not seem likely that the extra refinement needed to get a correction for levels with irrational term values can be derived theoretically by treating the polarization energy linearly.

* Now at the National Bureau of Standards, Washington, D. C.
[1] R. E. Trees, Phys. Rev. **83**, 756 (1951).
[2] R. E. Trees, Phys. Rev. **84**, 1089 (1951).
[3] E. Wigner, Phys. Rev. **46**, 1002 (1934); W. Macke, Z. Naturforsch. **5a**, No. 4, 192 (1950).
[4] R. E. Trees, Phys. Rev. **82**, 683 (1951).
[5] Values of the parameters are $A=19{,}969$, $B=953$, $C=3652$, $\alpha=67$.
[6] G. Racah, Phys. Rev. **85**, 381 (1951).
[7] The seniority number is introduced by G. Racah, Phys. Rev. **63**, 367 (1943).
[8] For instance, a $d^2\,{}^1H$ term has the energy $(A-17B+4C)$; with an $\alpha L(L+1)$ correction this becomes

$$A-17B+4C+30\alpha=(A-36\alpha)-17(B-2\alpha)+4(C+8\alpha).$$

The four parameters A, B, C, α are thus replaced by the three parameters $A-36\alpha$, $B-2\alpha$, $C+8\alpha$. This same substitution applies in all six levels.
[9] Values of the parameters are $A=18{,}400$, $B=853$, $C=4116$.

The Magnetic Threshold Field Curve of a Superconductor

R. P. Hudson
Cryogenics Section, National Bureau of Standards, Washington, D. C.
(Received November 15, 1951)

A number of authors[1] have considered the relation between the shape of the critical field *vs* temperature curve of a superconductor and the temperature dependence of the specific heat in both normal and superconducting states, from somewhat different points of view. In general, however, use is made of the Gorter-Casimir thermodynamic formulas[2] together with the assumption of a parabolic shape for the H_c *vs* T curve to obtain the familiar expression

$$\gamma=VH_0^2/2\pi T_0^2, \tag{1}$$

where γ is the "Sommerfeld term" (for the normal state), H_0 is the value of H_c at the absolute zero, V is the atomic volume, and T_0 is the zero-field transition temperature.

The complete expression one obtains is

$$\Delta C=C_s-C_n=-(VH_0^2/2\pi T_0^2)T+(3VH_0^2/2\pi T_0^4)T^3 \tag{2}$$

and on the assumption that C_n contains only a lattice term plus the Sommerfeld term, while C_s contains the same lattice term plus a T^3-term for the electron assembly, one derives relation (1) above; i.e., the linear term in (2) is identified with $-\gamma T$. We then have

$$\Delta C=-\gamma T+(3\gamma/T_0^2)T^3. \tag{3}$$

Furthermore, the jump in the specific heat at the zero-field transition temperature is given by

$$[\Delta C]_{T_0}=VH_0^2/\pi T_0=2\gamma T_0. \tag{4}$$

Miller[3] has reviewed this derivation [there are, incidentally, numerical errors in his equations, which are equivalent to (2) and (4) above] and remarks that it can be shown conversely, on the assumptions as to the expressions for C_s and C_n cited above, that the parabolic case necessarily follows. More recently Burns[4] restated this latter formulation in detail but there are, unfortunately, certain algebraic errors in his paper which should perhaps be pointed out. These appear to arise from an incorrect analysis of the general equation

$$H_c^2=H_0^2-(4\pi\gamma/V)T_0^2+(2\pi K/3V)T_0^4, \tag{5}$$

[Eq. (6) in Burns' paper], which is obtained by integration of the Gorter-Casimir expression and insertion of appropriate boundary conditions. Here,

$$H_0^2=(4\pi\gamma/V)T_0^2-(2\pi K/3V)T_0^4. \tag{6}$$

After solving (5) for the case $H_c=0$ we obtain

$$\frac{2\pi K}{3V}T_c^4-\frac{4\pi\gamma}{V}T_c^2+\frac{4\pi\gamma}{V}T_0^2-\frac{2\pi K}{3V}T_0^4=0,$$

if we define T_c as the value of T at which $H_c=0$, whence

$$[T_c^2-T_0^2][2\gamma-\tfrac{1}{3}K(T_c^2+T_0^2)]=0, \tag{7}$$

and the roots are given by

$$T_{c1}^2=T_0^2 \quad \text{and} \quad T_{c2}^2=6\gamma/K-T_0^2. \tag{8}$$

We will consider only the positive roots of Eq. (7) in the following; there are then, in general, two possible values for T_c. The root $T_{c1}=T_0$ must appear, of course, since that has been inserted by the stipulated boundary condition. (One must, incidentally, require that T_0 be real for there to be any meaning at all to the analysis!) There will also be a second root unless $T_0^2=3\gamma/K$, when the two roots coincide. From the algebraic point of view, the critical value of T_0^2 is $6\gamma/K$—not $3\gamma/K$ as given by Burns. If $T_0^2<6\gamma/K$ there will be two real roots although in between these H_c becomes imaginary; if $T_0^2>6\gamma/K$ the second root will be imaginary. Furthermore, in the latter case H_0 is also imaginary since H_0^2 is then negative.

For "acceptable" behavior of the H_c *vs* T curve, however, we must demand that it cross the T-axis at one point only, i.e., that $T_0^2=3\gamma/K$. This same conclusion is arrived at by Burns although his statement (8) is in error. [Here K is identical with $3\gamma/T_0^2$ in Eq. (3).]

To correct a typographical error in Eq. (14) of Burns' letter, the terms $(3\gamma/K)$ should be in brackets; it is then seen that this equation is identical with the more familiar expression (1) above. It should be emphasized that for the H_c *vs* T curve to be parabolic the essential condition is that ΔC should contain only terms in T and T^3. This is a somewhat wider restriction than that stated by Burns, *viz.* that the parabolic relation is found for elements whose electronic specific heat varies as T.

[1] See for instance, J. A. Kok, Leiden Commun. Suppl. 77a; Daunt, Horseman, and Mendelssohn, Phil. Mag. 27, 754 (1939).
[2] C. J. Gorter and H. Casimir, Physica 1, 306 (1934).
[3] A. R. Miller, Australian J. Sci. 10, 6, 172 (1948).
[4] G. Preston Burns, Phys. Rev. 76, 999 (1949).

Gravitational Deflection of Photons with Nonvanishing Rest Mass

ALBERT D. WHEELON
Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received November 28, 1951)

THE general theory of relativity relies on three experimental results for its validation: The precession of the perihelion of mercury, the gravitational shift of the spectrum, and the deflection of light rays as they pass the sun's rim. While the first two seem to be in good order, the initial confirmation[1] of the last has been seriously challenged by the Potsdam expedition[2] and others. Using measurements of the 1929 eclipse, they determined the angular deflection of light rays to be 2.10 ± 0.15 sec of arc, whereas the highest possible theoretical result is 1.75 sec of arc.

To account for the discrepancy, it is here assumed that the photons have a finite (though small) rest mass. The photons are then supposed to move in a plane under the action of the Schwarzschild field generated by the sun. The equations of motion may be integrated at once[3] to give

$$\left[\frac{du}{d\phi}\right]^2=u_0^2-u^2+\frac{2km}{c^2}\cdot\left[\frac{u-u_0}{p}+u^2-u_0^2\right], \tag{1}$$

in which we write $u=1/r$ and take u_0 as the maximum value of $u(\varphi)$ (i.e., $u_0=1/R$). We have subtracted the corresponding equation for u_0 and have eliminated the proper time by the first integral

$$r^2d\phi/ds=p. \tag{2}$$

In the usual discussion, the light rays are assumed to follow geodesics ($ds=0$), which in turn forces one to take p infinite. The calculation of the shift of the zeros of $u(\varphi)$ from the classical path may be carried out in an elegant fashion by a device of Bergmann.[4]

$$\delta\phi=4km/c^2R. \tag{3}$$

One need not, however, take $p=0$ to exploit the above method. The deviation in the case of $p\neq0$ can, in fact, be written in integral form.

$$\delta\phi=2\int_0^{1/R}du\left\{\left[(u_0^2-u^2)-\frac{2km}{c^2}\left(u_0^2-u^2+\frac{u_0-u}{p}\right)\right]^{-\frac{1}{2}}-[u_0^2-u^2]^{-\frac{1}{2}}\right\}. \tag{4}$$

If we multiply and divide by the small quantity $(2km/c^2)$, the integrand assumes the form $[f(x+\epsilon)-f(x)]\epsilon^{-1}$. We may approximate this by the usual derivative so as to obtain

$$\delta\phi=\frac{2km}{c^2}\int_0^{1/R}du\frac{u^2-u_0^2+\dfrac{u-u_0}{p^2}}{[u_0^2-u^2]^{\frac{1}{2}}}. \tag{5}$$

The integration is readily performed, and yields

$$\delta\phi=(4km/c^2R)[1+(R^2/2p^2)]. \tag{6}$$

The "proper angular momentum" p may now be determined by considering the photon at very large distances, and applying the special relativistic approximation.

$$r^2\frac{d\phi}{ds}=p=\frac{Rv_\infty}{c[1-v_\infty^2/c^2]^{\frac{1}{2}}}.$$

While the kinematical justification of this step is somewhat doubtful, it does not seem unduly bold to put $v_\infty=c$. Noting

$$h\nu=m_0c^2[1-v_\infty^2/c^2]^{-\frac{1}{2}},$$

one then finds

$$\delta\phi=(4km/c^2R)[1+\tfrac{1}{2}(m_0c^2/h\nu)^2]. \tag{7}$$

We now see that the effect of assigning a rest mass to the photon does indeed increase the predictable deflection, the additional term depending only on the rest mass and frequency of the light itself. In order that the additional term be of the proper magnitude to account for the larger observed values, however, it is necessary to assume

$$m_0c^2\simeq0.6h\nu.$$

For visible light ($\lambda=5000$A), we find $m_0\sim4\times10^{-54}$ g. This property should manifest itself in relatively simple laboratory experiments (for instance, the group velocity of photons should be $0.8c$), and since there is evidence to the contrary, we must abandon a finite photon mass as an explanation of the present

disagreement. One observes from this calculation, however, that frequency-dependent deflections are readily obtained from elementary combinations of familiar field relations. It is suggested that an experimental determination of the variation of displacements over available wavelengths, should prove helpful in understanding the gravitational field and its interaction with photon fields.

[1] See Campbell and Trempler, Lick Observatory Bulletin 11, 41 (1925); 13, 130 (1928).
[2] Freundlich, v. Kluber, and Brunn, "Ergebnisse der Potsdamer Expedition zur Beobachtung der Sonnenfinsternis von 1929, Mai 9, in Takengon (Nordsumatra)," Z. Astrophys. 3, 171 (1931).
[3] P. G. Bergmann, *Introduction to the Theory of Relativity* (Prentice-Hall, Inc., New York, 1942), p. 213 ff.
[4] See reference 3, p. 220.

The Effect of Neutron Irradiation on Metallic Diffusion

T. H. Blewitt and R. R. Coltman
Oak Ridge National Laboratory, Oak Ridge, Tennessee
(Received November 23, 1951)

THE study of the interaction of high energy neutrons with solids has been studied theoretically by Seitz[1] and observed experimentally by several investigators.[2-4] These investigations indicate that a fast neutron knocks-on an atom, displacing it from its lattice position. This displaced atom has sufficient energy to displace further atoms and hence a localized region with a large number of Frenkel defects should result. If Frenkel defects play a predominate role in the atomic diffusion of metals, neutron bombardment should greatly increase the rate of microdiffusion. On the other hand, macrodiffusion, that is the atomic diffusion over large distances, will be unaffected since the mean path for recombination of the defects will be small. An appropriate material for the study of this phenomenon is disordered Cu_3Au alloy, where every atomic motion will tend towards ordering.

The following experiment was performed. A sample of Cu_3Au was disordered, sealed in vacuum, and the electrical resistance measured at 200°C where the relaxation time is of the order of 10^5 hours.[5] The sample was then placed in an ORNL graphite reactor at a location where the fast neutron flux was of the order of 1×10^{12} neutrons per cm^2 per sec, with the temperature being maintained at 200°C. The electrical resistance was subsequently measured as a function of time and the results are shown in Fig. 1. A constant neutron flux was assumed. It is particularly interesting to note that a lag of 12 hours was observed when the pile was shut down (see Fig. 2) indicating that the relaxation time for ordering had decreased to about 12 hours at 200°C.

From these results and the room temperature bombardment of ordered Cu_3Au, an estimate can be made of the mean path for recombination of the Frenkel defects. The experiments on ordered Cu_3Au indicate that about 10^4 atoms are affected per atom

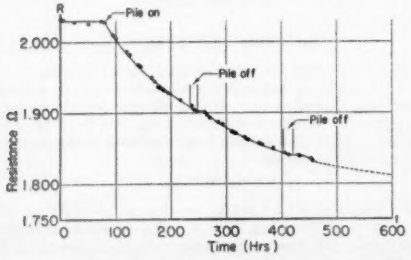


FIG. 1. Resistance of disordered Cu_3Au *vs* time. Temperature = 200°C. Fast neutron flux = 1×10^{13} n/cm²/sec.

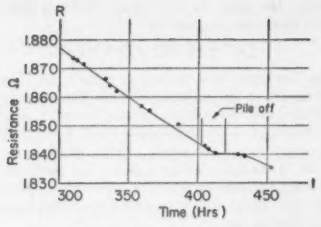


FIG. 2. Resistance of disordered Cu_3Au *vs* time. Temperature = 200°C. Fast neutron flux = 1×10^{13} n/cm²/sec.

knocked-on by a neutron.[3] From the data of Fig. 1 saturation should occur after about a month's exposure which is equivalent to 1×10^{18} neutrons per cm^2. Assuming the fast neutron scattering cross section to be 3 barns, the number of primary knock-ons is about 1×10^{17} per cm^3. If 10^4 atoms or regions about 40A in diameter are affected per primary knock-on, saturation should occur in about 100 months. Since the saturation appears to be reached in one month, then 10^3 atoms, or regions about 100A in diameter, should be affected, indicating that defects have a small but an appreciable mean path before annihilation.

The foregoing deduction can be verified by an experimental measurement of the size of the ordered domains. Since the order is not a single valued function of the resistance, it is not possible to deduce the degree of order from these measurements. It seems logical, however, to assume that the antiphase ordered regions have an equilibrium degree of order. If this assumption is made the resistance will indicate an average domain size of about 65A. X-ray studies are planned to determine the size of the domains and verify the above calculations.

If, as the theoretical treatments indicate, Frenkel defects are formed by fast neutron bombardment, then the above data offers strong evidence that an increase in the number of Frenkel defects will substantially increase the diffusion rate in metals. Further evidence that Frenkel defects play a role in metallic diffusion arise from measurements which indicate that the activation energy for the annealing of neutron radiation effects in copper is in close agreement with the activation energy for self-diffusion in copper.[6] Further studies on the radiation of disordered Cu_3Au are in progress and will be reported later.

[1] F. Seitz, Disc. Faraday Soc. No. 5, 271 (1949).
[2] S. Siegel, Phys. Rev. 75, 1823 (1949).
[3] D. S. Billington and S. Siegel, Metals Progress 847 (December, 1950).
[4] J. H. Crawford *et al.*, Phys. Rev. 83, 312 (1951).
[5] F. Sykes and H. Evans, J. Inst. Metals 58, 255 (1936).
[6] To be published.

Properties of Powdered $BaTiO_3$

V. E. Derr and M. D. Earle
Radiation Laboratory, Johns Hopkins University, Baltimore, Maryland
(Received November 19, 1951)

RECENTLY much interest has been shown in the properties of ceramic barium titanate. In this laboratory an effect has been found to exist in the powdered form of $BaTiO_3$ which is present to only a negligible extent in the ceramic material. These experiments were conducted on capacitors whose plates were approximately 35 mm in diameter and 0.5 mm apart, and whose dielectric consisted of powdered $BaTiO_3$. The first experiments were conducted on capacitors in which the powder had been packed by hand pressure only. Later the experiments were repeated, using a hydraulic press to pack the powder to a much higher density than could be done by hand.

It was found that if these capacitors were initially charged (the charging source was a 45-volt battery) and then short circuited for

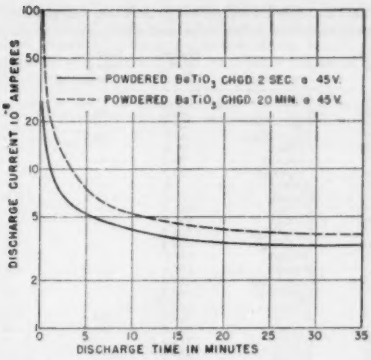


FIG. 1. Discharge characteristics of capacitor with powdered $BaTiO_3$ as dielectric.

a few seconds, an appreciable voltage was present after the short circuit had been removed. This voltage gradually decreases with time, dropping from a value of about 1 volt when the short circuit was removed to a value of about 0.1 volt at the end of 40 minutes. To avoid the possibility of error caused by the vacuum tube voltmeter used for measuring the voltage, a 1047-ohm galvanometer in series with a 300,000-ohm resistor was connected across the capacitor and the current was observed as a function of time. If the capacitor was short circuited for a few seconds during this time, upon removal of the short circuit the current was found to return to approximately its previous value.

These capacitors have a relatively low value of resistance. The resistance as measured with a battery-operated ohmmeter appears to increase from a value of approximately 100,000 ohms to approximately one megohm after two or three minutes. The capacity of these capacitors at 1000 cycles was approximately 10,000 $\mu\mu f$. The charge obtained from these capacitors appears to be greater than that placed on them if the formula $Q=CV$ is used to determine the charge. The use of this formula does not seem to be permitted in this case, however, since the amount of current that can be obtained from the capacitor is much reduced if it is charged from a mica capacitor so that the charge of the barium titanate capacitor is about equal to the value computed by the formula $Q=CV$. Much more charge than that indicated by the formula seems to be placed on the $BaTiO_3$ capacitor. However, the effect of charging the capacitor for a long period of time before disconnecting from the charging source does not seem to have a very great effect, as shown by two curves of current plotted against time (see Fig. 1). The charging times were 2 seconds and 20 minutes. Each of these periods is very long compared to the time constant of the charging circuit which is approximately 10^{-6} sec. A semi-log plot is used to show these data. The curvature of the lines shows clearly that the change of current with time differs from that of a conventional condenser. The results were changed only slightly when different metals were used as condenser plates.

The Molecular Spectrum of Iodine Excited in the Presence of an Inert Gas

K. WIELAND*
University of Zürich, Zürich, Switzerland
AND
JÖRG WASER
Rice Institute, Houston, Texas
(Received December 4, 1951)

RANK'S recently corrected vibrational formula[1] for the ground state of diatomic iodine (referred to below as II) differs considerably for large v''-values from his earlier formula[2] (to be called I hereafter). This makes necessary a revision of the vibrational analysis of the band systems[2] $D\leftrightarrow X$, $C\rightarrow X$, and $F\rightarrow X$, which, if excited in fluorescence[4] or in emission[5,6] in the presence of an inert gas, involve high-lying vibrational levels $G(v'')$ of the ground state X. System $E\rightarrow B$, which similarly is obtained in the presence of an inert gas,[5] is of course not affected by this revision. The results of our revision for $D\leftrightarrow X$ and $F\rightarrow X$ are presented in Table I, whereas $C\rightarrow X$ will be discussed at the end of this letter.

The revised constants of the vibrational levels of the state F are not significantly different from the ones obtained[2] with Rank's formula I. At the highest v''-values, however, which now run up to 75, our new vibrational analysis differs considerably from the one given by Venkateswarlu. In addition, the fit to the observed data obtained with Rank's earlier formula I is not nearly as good as the one we obtained using Rank's newer formula II, with the proviso that the constant $t_{\omega\omega_e}=1.60\times10^{-7}$ given by Rank[2] be replaced by 1.45×10^{-7} cm^{-1}. This changes $z_{\omega\omega_e}$ from 5.25×10^{-6} to 5.29×10^{-6}, $z_{\omega\omega_e}=5.65\times10^{-6}$ remaining unchanged.

We find Venkateswarlu's analysis of system $D\leftrightarrow X$ untenable, since his high $G''(v)$ values running up to $G''(76)=11{,}500$ cm^{-1} are hardly compatible with the fact that the bands of system $D\leftrightarrow X$ can also be observed in absorption by iodine vapor in a 26-cm tube at 1000°C and 1 atmos.[7] Furthermore, the intensity distribution expected on the basis of his assignment is quite different from the one observed. Indeed, our revision using formula II, including the small modification of the constant $t_{\omega\omega_e}$ mentioned above, reveals that the previous analysis by Elliott[4] was very nearly correct. Our revised formula lowers Elliott's v''

TABLE I. Revised[a] vibrational constants of I_2.

State[b]	T_e	ω_e	$z_{\omega\omega_e}$	$\gamma_{\omega\omega_e}$	$z_{\omega\omega_e}$	$t_{\omega\omega_e}$	Obs. trans.	$\nu_{0,0}$
$X^1\Sigma_g^+$	0	214.248	0.6074	−0.00130	-5.529×10^{-6} [c]	-1.45×10^{-7} [c]		
$D^1\Sigma_u^+$	33,346.6	104.0	0.2				$D\leftrightarrow X$	33,291.6
$F(^3\Pi_u)$	47,208.25	96.7[d]	0.49				$F\rightarrow X$	47,147.5

[a] States $A(^3\Pi_{1,u})$, $B(^3\Pi_{0,u}^+)$, and $E(^1\Sigma_g^+)$ are as given by Herzberg. (Through an error, Herzberg gives T_e of state E as 41,407 instead of 41,411 cm^{-1}.) State $C(^1\Sigma_{1,u}^+)$, called E by Venkateswarlu, needs revision to be adapted to Rank's formula II. However, the present interpretation of the observed system is not necessarily correct (see reference 6 in text). The vibrational analysis and interpretation of the many systems with $\nu>50{,}000$ cm^{-1} must still be regarded as uncertain.
[b] Designation used by Herzberg.
[c] Rank's values are: $z_{\omega\omega_e}=-5.25\times10^{-6}$ and $t_{\omega\omega_e}=-1.60\times10^{-7}$ cm^{-1}.
[d] Herzberg erroneously gives for ω_e the $\omega_e x_e$ value (96.21 cm^{-1}).

numbering by 2 units, thus reducing the highest observed $G''(v)$ value to $G''(26)=5115$ cm^{-1}. It includes, in the region λ3454–3237A, all of the 37 fluorescence bands observed by Elliott as well as nearly 90 additional emission bands obtained by Venkateswarlu, with the exception of eight presumably rather weak bands none of which were observed in fluorescence or absorption. Below λ3237A, however, more and more bands appear which do not fit into our Deslandres table, and are assumed to belong to another system.

The vibrational analysis of system $C \rightarrow X$, given by Venkateswarlu[6] using Rank's formula I, need not undergo a very significant change to be brought in accord with Rank's formula II. We should like, however, to emphasize the fact mentioned by Elliott[4] that this small system centered at λ2767A (together with two other diffuse bands at 2829 and 2878A) can be excited with active nitrogen but not in fluorescence with light of λ1854A. The most likely explanation[8] of this fact is that the upper state lies higher than 54,000 cm^{-1} and combines with a state lying above the ground state, for instance with the term $B^3\Pi_{0u}{}^+$. For this reason no constants for state C were included in the table.

A detailed paper containing the observed data of the systems $F \rightarrow X$ and $E \rightarrow B$ will appear later in the *Helvetica Physica Acta*.

* Now at the Department of Physics, University of California, Berkeley, California.

[1] D. H. Rank and W. M. Baldwin, J. Chem. Phys. **19**, 1210 (1951).
[2] D. H. Rank, J. Opt. Soc. Am. **36**, 239 (1946).
[3] This designation is the one used by G. Herzberg, *Spectra of Diatomic Molecules* (D. Van Nostrand Company, Inc., New York, 1950), p. 541, and differs from the one used by Venkateswarlu.
[4] A. Elliott, Proc. Roy. Soc. (London) **A174**, 273 (1940).
[5] J. Waser and K. Wieland, Nature **160**, 643 (1947).
[6] P. Venkateswarlu, Phys. Rev. **81**, 821 (1951).
[7] E. Skorko, Nature **131**, 366 (1933) and Acta Phys. Polon. **3**, 191 (1934).
[8] It is difficult to check this suggestion from the data published by Venkateswarlu without having a reproduction of the spectrum showing the intensity distribution of the bands.

The Radioactive Decay of Hg^{197} and Hg^{203}†

J. M. Cork, D. W. Martin, J. M. LeBlanc, and C. E. Branyan
Department of Physics, University of Michigan, Ann Arbor, Michigan
(Received November 8, 1951)

ON bombarding gold with protons or deuterons a radioactive isotope of mercury-197 had been found[1,2] which decayed showing half-lives of about 25 hr and 65 hr. It had been assumed that 'K'-electron capture occurred in each activity, with no isomeric transition in mercury. All of the observed gamma-emissions were supposed to follow in the resulting gold nuclei. Gamma-energies of 77, 135, 165, and 273 kev were reported. An excited level in gold-197 with a half-life of 7.5 sec, first found by Wiedenbeck by gamma-excitation, was again observed in the mercury decay. Coincidence experiments had shown[3,4] the 135 and 165-kev gammas to be in cascade with a positive angular correlation existing between them.

It is now possible by spectrometric studies of sources irradiated in the Argonne heavy-water reactor to express more exactly the energies of the gamma-rays and to show the presence of one gamma-ray not previously observed. Moreover, the conversion electrons of two strong gamma-rays, 134.3 and 165.4 kev, are found to satisfy the work functions characteristic of mercury rather than of gold as had been reported, and are hence associated with an isomeric transition before K-capture. A continued sequence of spectrograms was taken to show precisely which electron conversion lines died out with a half-life of 25 hours and which persisted with a longer half-life. Only those electron lines associated with the 134.3- and the 165.4-kev gammas decayed with the 25-hr half-life. The energies of the electron lines, together with their interpretations, are shown collectively in Table I. Arbitrary numbers in the order of increasing energy are assigned to the gamma rays and shown as superscripts in column 2.

TABLE I. Conversion electron energies and their interpretation.

Electron energy (kev)	Interpretation	Energy sum (kev)	Gamma energy (kev)	Ratio K/L
50.9	K^3(Hg)	134.1	···	···
52.0	$A(\alpha_2-L_I)$	66.9	66.9	···
54.1	$A(\alpha_1-L_I)$	68.7	···	···
56.5	$A(\alpha_1-L_{III})$	68.7	68.7	···
63.3	$L_I{}^4$(Au)	77.6	···	···
63.9	$L_{II}{}^4$(Au)	77.6	···	···
65.6	$L_{III}{}^4$(Au)	77.5	···	···
74.2	M^4(Au)	77.6	···	···
76.8	N^4(Au)	77.6	77.6	$L_I/M=4$
82.3	K^3(Hg)	165.4	···	···
110.8	K^4(Au)	191.5	···	···
119.7	$L_{I,II}{}^3$(Hg)	134.5	···	···
122.0	$L_{III}{}^3$(Hg)	134.3	···	···
131.2	M^3(Hg)	134.9	···	···
133.2	N^3(Hg)	134.0	134.3	<0.1
150.6	$L_{I,II}{}^5$(Hg)	165.4	···	···
152.7	$L_{III}{}^5$(Hg)	165.3	···	···
161.6	M^5(Hg)	165.5	···	···
164.2	N^5(Hg)	165.0	165.4	~0.25
177.1	$L_I{}^4$(Au)	191.2	191.4	~9
194.0	K^4(Tl)	279.5	···	···
264.2	$L_I{}^4$(Tl)	279.5	···	···
275.6	M_I(Tl)	279.3	279.5	~10

Mercury-203 was also produced in the reactor from an enriched (97.3 percent) isotope of mercury-202. On following its decay over several octaves a half-life of 47.9±0.2 days is observed. A single gamma is emitted whose energy is found to be 279.5 kev. This activity was always present in the shorter-lived mercury specimens. Its strong electron lines in thallium are so situated as to almost exactly mask any weaker conversion lines because of a transition in gold of 273 kev, had such electron lines been present in the 25-hour activity as reported by Huber *et al.*

A half-life of 66.4 hours is found for the intermediate period, from observations extending over a month, with corrections for the presence of the 47.9-day activity. No attempt was made to check the reported half-life of 25 hr.

A proposed level scheme is shown in Fig. 1. The order of the 134.3–165.4 kev sequence is proposed from their relative K/L ratios. Deutsch had observed a delay of 1×10^{-8} sec in a coincidence study of the conversion electrons for the two gammas. The electron lines associated with the 77.6-kev gamma-ray are very strong both by conversion and photoelectrically in lead. The K series α_1, α_2 x-ray lines of gold are fairly strong as shown by

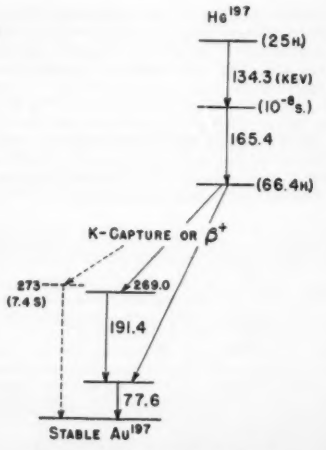


FIG. 1. Proposed nuclear level scheme.

Auger lines and by photoelectrons from a lead radiator. The reported 273-kev transition with its 7.5-sec half-life, although not observed, is included with dotted lines as shown. It could be that this gamma-ray is identical with the 269-kev transition allowed in the level scheme.

Note added in proof.—A more recent paper[6] has just come to our attention in which the authors have arrived at conclusions similar to those here expressed.

† This investigation was made possible by the joint support of the AEC and ONR.

[1] G. Friedlander and C. Wu, Phys. Rev. **60**, 747 (1941); **63**, 227 (1943).

[2] Frauenfelder, Gugelot, Huber, Medicus, Preiswerk, Scherrer, and Steffen, Helv. Phys. Acta **20**, 238 (1947); Huber, Steffen, and Humbel, Helv. Phys. Acta **21**, 192 (1948).

[3] Frauenfelder, Walter, and Zünti, Phys. Rev. **77**, 557 (1950).

[4] M. Deutsch and W. Wright, Phys. Rev. **77**, 139 (1950); F. McGowan, Phys. Rev. **77**, 138 (1950).

[5] Huber, Humbel, Schneider, de Shalit, and Zünti, Helv. Phys. Acta **24**, 127 (1951).

Proceedings of the American Physical Society

MINUTES OF THE MEETING OF THE DIVISION OF ELECTRON PHYSICS AT THE NATIONAL BUREAU OF STANDARDS, WASHINGTON, D. C., NOVEMBER 1–3, 1951

Symposium on Electron Emission Phenomena in Commemoration of the Bureau's 50th Anniversary

A1. Address of Welcome. W. R. BRODE, Assistant Director, *National Bureau of Standards.*

A2. Atomistic Theory of the Metallic Surface. CONYERS HERRING, *Bell Telephone Laboratories.*

A3. Electrical Properties of Porous Semiconductors.* E. B. HENSLEY, *University of Missouri (now at Research Laboratory of Electronics, M.I.T.).*—Currently in the literature there are two theories concerning the conduction mechanism in oxide coated cathodes. One regards the crystals of the porous aggregate as simple *n*-type semiconductors with the conduction current passing electronically through the crystals. The other maintains that the conduction current at the higher operating temperatures is carried by the electron gas in the pores of the aggregate. An analysis of this second hypothesis will be discussed pointing out the conditions under which pore conduction will occur. The results of this analysis show that while the temperature dependence of the conductivity of an oxide cathode is compatible with the pore conduction theory, the ratio of the conductivity to the thermionic emission as reported in the literature is somewhat too large. An aggregate in which there are two parallel mechanisms for conduction will, in the presence of thermal gradients, possess two sources of thermoelectric emf. A simple theory for the thermoelectric power of the electron gas in the pores of a porous semiconductor is developed and the manner in which it combines with the normal thermoelectric power of the crystals will be analyzed. In most respects the results are in good agreement with the experimental data obtained by Young. However, the essential feature in this agreement appears to be the presence of two parallel mechanisms.

* Supported in part by the ONR.

A4. Electrical Conductivity and Thermoelectric Power of (BaSr)O and BaO.* J. R. YOUNG, *University of Missouri (now at General Electric Research Laboratory).*—A study has been made of the thermoelectric power, the electrical conductivity, and the thermionic emission of (BaSr)O and BaO at different states of activation and on different base metals over a temperature range of 1100°K to 300°K. A plot of the logarithm of the conductivity *vs* reciprocal temperature showed two straight line regions. The temperature dependence of the thermionic emission was found to be the same as the temperature dependence of the conductivity in the temperature range 700°K to 1100°K. The slope of a plot of the thermoelectric power *vs* reciprocal temperature was also similar to the thermionic work function and the conductivity activation energy in this temperature range. The thermoelectric power varied with temperature from a value of about 2.0 mv/° at 1100°K, increasing to about 2.5 mv/° at 800°K, reducing to between 1.0 and 0.5 mv/° at about 500°K, and remaining practically constant or increasing slowly at still lower temperatures. In all thermoelectric power measurements the hot junction was positive indicating negative charge carriers. The simple semiconductor theory does not adequately account for these results which seem to be in good agreement with the pore conduction hypothesis.

* Supported in part by the ONR.

A5. Application of Nijboer's Theory to the Semiconducting Emitters and Its Modification. T. ARIZUMI AND S. NARITA, *Kobe Kogyo Corporation.*—Applying Nijboer's[1] theory to the semiconducting emitter the emission formulas applicable to any activated state have been easily deduced but recent studies on the Hall coefficient and electrical conductivity seem to suggest the coexistence of both *n*- and *p*-type conduction. Our measurements on the conductivity of the oxide coated cathode, in which pure MgO was used as the base, support this suggestion and moreover, recent experiments of Ishikawa and others[2] have revealed that the conductivity increased gradually and at the same time, the Hall coefficient changed its sign at higher oxygen pressures of about 10^{-1} to 100 mm Hg[3] Assuming that excess oxygen atoms are responsible for the formation of *p*-type level, some modification to the previous theory will be postulated. The obtained results are qualitatively in satisfactory agreement with the experiments on the conductivity and the thermionic work function.

[1] B. R. A. Nijboer, Proc. Phys. Soc. (London) **51**, 575 (1939).
[2] Y. Ishikawa and others, unpublished (oral presentation to Japan Electron Emission Committee).

B1. Chemical Reactions in Barium Oxide on Tungsten Emitters. R. C. HUGHES, P. P. COPPOLA, AND H. T. EVANS, *Philips Laboratories, Inc.*—When barium carbonate is heated in vacuum in contact with tungsten, the following reactions may occur in sequence as progressively higher temperatures are reached: (1) $WO_3+BaCO_3 \leftrightarrows BaWO_4+CO_2$; (2) $3\,BaCO_3+W \leftrightarrows Ba_3WO_6+3CO$; (3) $BaCO_3 \leftrightarrows BaO+CO_2$; (4) $6BaO+W \leftrightarrows Ba_3WO_6+3Ba$. At approximately 600°C, reaction (2) occurs with appreciable speed. Reaction (3) proceeds at 800–900°C. Reactions (2) and (3) may proceed simultaneously, preponderance of one over the other depending upon state of subdivision and intimacy of admixture of the reactants. Under favorable conditions reaction (2) may be driven to completion with almost total absence of reaction (3). Reaction (4) progresses with appreciable speed over approximately 1000°C.

Activated cathodes of BaO on W will normally have an interface of Ba_3WO_6 formed by reactions (2) or (4), or both. The end result of heating the cathode is the conversion of the entire thickness of the oxide coating to the basic tungstate. The resulting coating of Ba_3WO_6 on W is unreactive and nonvolatile at temperatures up to 1300°C, and provides no significant thermionic emission.

B2. The Production of Sr Metal in the Reduction of SrO by Methane. H. W. ALLISON AND GEORGE E. MOORE, *Bell Telephone Laboratories.*—The rate of this reaction is readily measurable at temperatures above about 1100°K by using radioactive isotopes. A pure platinum wire, serving as the mechanical support for radioactive SrO, can be heated to any temperature of interest but is itself chemically inert. The SrO can thus be maintained at a pyrometrically observed temperature in methane and the Sr metal, produced in the reaction, deposits on the bulb or on auxiliary filaments where it is measured by counting techniques. Present results indicate that the methane first dissociates into carbon and hydrogen, the carbon being primarily responsible for the reduction. SrO is reduced by methane at a rate considerably more rapid than

by tungsten,* but less rapidly than the theoretical maximum rate, which can be predicted from thermodynamics.

Atomic hydrogen should leave no contaminant, such as carbon, in the oxide; and thermodynamic calculations indicate that it should be a more powerful reducing agent than methane. It will be examined experimentally.

* G. E. Moore and H. W. Allison, J. Chem. Phys. 18, 1579 (1950).

B3. The BaSrO Cathode Supported on an Insulator. GEORGE E. MOORE AND H. W. ALLISON, *Bell Telephone Laboratories.*—Preliminary results were presented in 1949.* Much physical and chemical complexity inherent in practical (BaSr)O cathodes could be eliminated and the interpretation of experimental data thus simplified if one could investigate the (BaSr)O layer without its usual metallic support. This was accomplished by supporting the layer on MgO ceramic, a chemically inert insulator. Measurement of electron emission from the insulated (BaSr)O required microsecond pulse technique. Pyrometer temperatures of the translucent cathodes were corrected by determining radioactively the evaporation rate of BaO and SrO, which are known functions of true temperature. Electron emission from these nonmetallically supported (BaSr)O layers is at least 5 orders of magnitude greater than from the uncoated MgO, a result consistent with generally accepted semiconductor theory which assigns no fundamental importance to the metal support. However the emission is 2 to 3 orders of magnitude less than from practical cathodes and is not increased significantly by treatment in atomic hydrogen or methane. Evidence given in the preceding paper indicates that BaO should be reduced by such treatments, whose ineffectiveness is therefore difficult to explain in terms of theories which attribute the emission essentially to excess Ba.

* G. E. Moore and H. W. Allison, Phys. Rev. 77, 246 (1950).

B4. Diffusion Phenomena in the Oxide Cathode. E. S. RITTNER, *Philips Laboratories, Inc.*—Thermochemical calculations indicate that the equilibrium pressure of barium (and strontium) produced by reaction between the activator metal and (Ba-Sr)O is sufficiently high so that the rate of reaction is limited by the diffusion rate of the activator in the core nickel.[1] Similarly, the equilibrium pressure of the dissolved barium above the oxide is high enough so that the rate of barium depletion during the life of the cathode is limited by the rate of diffusion of barium out of the oxide. A theoretical survey of diffusion phenomena in the oxide cathode based upon recently determined diffusion coefficients[2,3] has yielded the following additional information. The activator-oxide reaction goes to completion in the very early stages of the activation schedule. The barium vapor so generated is transported throughout the porus oxide mass by means of Knudsen flow. During the activation period barium can penetrate into the oxide crystals for a distance of only a few microns. The subsequent slow diffusion of barium out of the oxide at operating temperatures accounts for the long life of the cathode.

[1] M.I.T. Conference Report on Physical Electronics, March 29, 1951, p. 24.
[2] G. F. Rouse and R. Forman, Phys. Rev. 82, 574 (1951); private communication.
[3] R. W. Redington, Phys. Rev. 82, 574 (1951).

B5. The Solubility and Diffusion Coefficient of Carbon in Nickel. J. J. LANDER, H. E. KERN, AND A. L. BEACH, *Bell Telephone Laboratories.*—The weight percent solubility of carbon in nickel varies from 0.041 at 780°C to 0.245 at 1030°C. These results were obtained from carbon saturated 499 nickel sheet analyzed according to the method developed by L. A. Wooten and W. G. Guldner.

Values obtained for the diffusion coefficient of carbon in nickel range from 4.0×10^{-9} cm² per sec at 727°C to 4.1×10^{-7} at 1020°C. These values were obtained by a novel method in which carbon coated on one side of a nickel disk is diffused through the disk to react with nickel oxide coated on the other. Rates of diffusion are observed as rates of evolution of the gaseous products of reaction. The initial, steady, and final states predicted by diffusion theory may be accurately observed if proper experimental conditions are satisfied. The results have been applied in an analysis of reactions taking place in oxide coated thermionic cathodes. It was found in experiments with nickel coated on one side with BaO and on the other with C that rates of reaction between C and BaO were nearly identical with those between NiO and C. Thus the rate limiting factor in the reactions is the diffusion of carbon through nickel.

B6. The Electronic Properties of Barium Sulfide.* W. GRATTIDGE† AND H. JOHN, *University of Missouri.*—Barium sulfide has been prepared as a cathode emitter and in the form of compressed pills. It has a work function of 2.1 ev with a wide variation possible in activation. The most active materials had an emission at temperatures of 900°K and greater, comparable to that from pure BaO. Pure nickel and platinum were used as base metals. An equimolar solid solution of BaS−S+S had a work function of 2.6 ev. The electrical conductivity of BaS depended on the degree of activation and the denseness of the specimen but was much less than for BaO. At 1000°K the specific conductivity of a compressed pill was 10^{-6} ohm^{-1} cm^{-1}. The thermoelectric power was constant with temperature, decreasing with increasing activity, the warmer electrode always positive. Pure BaS had a thermoelectric power of 2.5 mv per degree between 700 and 1000°C. The effect of added impurities of iron was studied. Mass-spectrometer studies indicated a vapor pressure of the sulfide about one hundred times less than that of the oxide. The latent heat of vaporization was 104 kcal/mole.

* Supported in part by the ONR.
† Now at Massachusetts Institute of Technology, Cambridge, Massachusetts.

C1. The Use of Radioactive Isotopes in the Study of the Evaporation from Thermionic Cathodes.* W. F. LEVERTON AND W. G. SHEPHERD, *University of Minnesota.*—The quantity of the alkaline earths transferred from oxide coated cathodes during processing and life has been measured. Radioactive tracer techniques are used. The isotopes Ba140, Sr89, and Ca45 are incorporated singly into carbonate cathode coating mixtures. The radioactive cathodes are used in diodes having a movable electrode which shields the anode from the cathode during all or any desired part of the tube processing schedule. At the conclusion of the test the diode is opened and the radioactivity on the anode, movable shield, and cathode is measured. This method has proven very useful in determining evaporation rates for Ba, Sr, and Ca as a function of cathode operating temperature, emission current, processing schedule, and life. Experiments are in progress to study the behavior of the double carbonate of Ba and Sr and the single carbonates of Ba, Sr, and Ca.

The tracer technique allows the detection and accurate measurements of quantities of cathode Ba, Sr, and Ca as small as 10^{-10} g.

* This work was supported by Signal Corps contract.

C2. Photoelectric Emission from BaO in an Accelerating Field. PAUL E. CARROLL* AND E. A. COOMES, *University of Notre Dame.*—A consideration of the field dependent photoelectric effect reveals a convenient method for determining work functions of photoelectric emitters. By extending Fowler's photoelectric theory for a lowering of the potential barrier at the surface of the emitter due to the Schottky effect, it can be shown that the current should be proportional to the

square root of the applied field for energies of input light a few tenths of an electron volt from the threshold. The slope of I *vs* $E^{\frac{1}{2}}$ is proportional to the work function. Nine independent determinations of the work function of a BaO cathode using this method are in agreement to within one percent. This method has the advantage over the use of Fowler plots in that a normalization to unit intensity of input light is not required. In general, the results on BaO are in qualitative agreement with the Guth-Mullin theory for photoemission in an accelerating field. Indications are that the true field at the BaO cathode is several times greater than the geometrically computed field. This would substantiate the thermionic Schottky emission data for alkaline earth oxides reported in the literature.

* Now at the Sylvania Electric Products, Inc., Kew Gardens, New York.

C3. Photoconductivity Study of Activation of Barium Oxide. H. B. DEVORE, *R.C.A. Laboratories.*—Photoconductivity studies have been made for barium oxide cathodes at various stages of thermionic activation, for the purpose of attempting to establish the donor energy levels. Changes occurring during activation are reflected in the development of the long wavelength tail of the photoconductivity curve. For states of low activity, this tail has a threshold in the neighborhood of 2.3 electron volts, and, with increasing thermionic activity, a second threshold at about 1.5 electron volts develops. The measurements are interpreted as indicating the development of donors at levels having these depths below the conduction band during the activation. Tentative explanations of the origin of these donors are advanced.

C4. The Luminescence of BaO.* VIRGIL L. STOUT, *University of Missouri (now at Stanford Research Institute).*—The luminescence of BaO has been studied in an endeavor to correlate its luminescent and thermionic properties. BaO samples were prepared by heating $BaCO_3$ in a vacuum. The carbonate was sprayed with a binder on two types of cathode base metals. These sample cathodes were mounted in a vacuum tube designed so that thermionic emission studies were possible at temperatures up to 1000°K and luminescent measurements could be made at temperatures down to 90°K. Measurements were made of thermionic emission and luminescent intensity distribution both as a function of the state of thermionic activation, and as a function of temperature. Results of these measurements have been interpreted in terms of a conventional model for BaO involving both donor and acceptor type impurities. Luminescence mechanisms have been proposed for the luminescent bands at 345 mμ and at 465 mμ which vary as a function of thermionic activation. The temperature dependence of the 345 mμ band was associated with the temperature dependence of the mean free path of a positive hole. From this association, one is able to determine the characteristic temperature of the fluorescent material.

* This work was supported in part by the ONR.

C5. On the *I–V* Characteristics of the "*L*-Cathode." E. S. RITTNER AND R. H. AHLERT, *Philips Laboratories, Inc.*—Measurements of the *I–V* characteristics of a planar diode incorporating an *L*-cathode have been made in the temperature range 1150–1400°K and in the voltage range 0–1500v. The data were taken with the use of a circuit* of the type previously employed by Loosjes and Vink,[1] which periodically discharges an *RC* network across the diode. In our work the time constant of the network was 60 μsec and the pulse repetition rate was 60 per second. The results are in accord with theoretical expectations in both the space charge and Schottky regions. Typical parameters obtained from a zero field Richardson plot are $\phi = 1.7$ ev, $A = 3.0$ amperes/cm^2 deg^2. Some observations regarding poisoning phenomena will also be reported.

* We are indebted to T. R. Kohler for the specific design employed.
[1] R. Loosjes and H. J. Vink, Philips Research Reports **2**, 190 (1947).

C6. Physical Processes in the "*L*" Cathode. D. L. SHAEFER AND J. E. WHITE, *General Electric Company.*—Experiments have been carried out to investigate further the nature of this emitter, whose mechanism has been discussed in a previous paper.[1] The pattern described by duPre and Rittner agrees in general with the results of the present work. Evidence has been accumulated to show that barium is delivered to the surface of such a cathode by surface diffusion and that the streaming of vapor from the pores is negligible.

In looking toward means of combating excessive barium evaporation, it was important to establish whether less or more than a monolayer of barium resides on the surface during operation. Two independent approaches have established that less than a monolayer exists in equilibrium under normal operating conditions. The evaporation rate of active material has been measured as a function of temperature and is compared with values for other commonly used thermionic emitters. The evaporant has been identified as barium metal.

[1] F. K. duPre and E. S. Rittner, Phys. Rev. **82**, 573 (1951).

W. G. SHEPHERD presiding

ThO_2, Secondary Emission

D1. Electrolysis of Thorium Dioxide. I. Oxygen Evolution from Operating Cathodes. EDWARD SHAPIRO, *Bartol Research Foundation.*

D2. Electrolysis of Thorium Dioxide. II. Observations with Homogeneous Crystalline Specimens. W. E. DANFORTH, *Bartol Research Foundation.*

D3. The Angular Distribution of Secondary Electrons from Nickel. J. L. H. JONKER (PRESENTED BY F. K. DUPRE), *Philips Gloeielampenfabrieken.*—The common equipment for measuring secondary emission (disk-shaped or spherical electrode within a sphere) is not suitable for obtaining data about the angular distribution of secondary electrons. An electrode system with two concentric spheres was constructed in order to obtain a really radial retarding field with which the behavior of the secondary electrons with different velocities could be studied. The angular distribution of three groups of secondary electrons, slow genuine secondary electrons, secondary electrons with moderate velocities, and rapid reflected electrons, was measured as a function of the angle of incidence and of the energy of the primary electrons. The experimental results may be understood assuming Whiddington's law concerning the energy loss of electrons that penetrate into a solid substance, an isotropic distribution of the secondaries within the substance and an exponential absorption on their way out.

D4. Secondary Electron Emission from Germanium. J. B. JOHNSON AND K. G. MCKAY, *Bell Telephone Laboratories.*—Secondary emission by electron bombardment has been measured for single crystals of Ge having a $p-n$ junction. The highly purified Ge was in one case doped with Ga for the p-type end, and in the other case heavily doped with Sb for the n-type end and with additional Ga for the p-type end. The carrier densities far from the junctions ranged from 2×10^{14}/cm^3 to 10^{17}/cm^3. The crystals were cut and polished along a (100) plane, chemically etched to give a clean surface, and given prolonged heat treatment up to 700°C in high vacuum during preparation of the tube. The two crystals, or the n- and p-halves of the same crystal face with the same treatment, show no difference in yield as large as 2 percent in spite of the difference in calculated surface fields. The yield curve in the range $V_p = 50$–5000 volts is very similar to that of Keller and Burgess,* having a $\delta_{max} = 1.1_5$ at 600 volts. Normalized energy distribution of secondaries shows no change over

V_p=500–5000 volts, with most probable energy at 1.5–2.0 volts, and is the same on the p and n sides of both crystals. At 600°C δ is about 2 percent lower than at room temperature.

[9] R. L. Koller and J. S. Burgess, Phys. Rev. **70**, 571 (1946).

D5. Decay Time of Secondary Electron Emission. R. R. Law, *R.C.A. Laboratories.*—Tests on a series of developmental tubes using multi-stage electrostatic electron multipliers show unexpectedly poor performance at high frequencies. The efficiency of a Class-*C* power amplifier with a five-stage multiplier is found to decrease with frequency from 0.7 at 50 mc to 0.2 at 450 mc. This result is obtained with both large-scale and small-scale tubes (over-all transit times of 2×10^{-9} and 5×10^{-10} second, respectively) with different input systems and with different dynode surfaces. The high frequency performance is substantially independent of voltage, temperature, current density, modulation index, or dynode material. The dynode materials so far tested include MgO, BaO, BeO, and gold.

These results may be explained by assuming that each stage of the electron multiplier exhibits a simple expotential decay characteristic with a time constant of 3×10^{-10} second. In view of the reproducibility of the effect with different tube structures, this time constant cannot reasonably be ascribed to circuit difficulties or known defects of electron multipliers. It is suggested that the secondary electron emission process itself introduces this characteristic; that is, that the fundamental process of secondary electron emission has a simple exponential decay characteristic with a time constant of about 3×10^{-10} second.

D6. The Wave-Mechanical Foundation of Baroody's Theory of Secondary Emission.* A. J. Dekker† and A. van der Ziel, *University of Minnesota.*—Baroody's semiclassical theory[1] of secondary emission can be derived with the help of wave mechanics. It can be shown that those collision processes between primary electrons and lattice electrons in which the lattice does not take up momentum give Baroody's results. The collision processes in which the lattice does take up momentum have been studies by Wooldridge[2] and have been applied to the problem of secondary emission. A closer inspection of the formulas shows, however, that for energies below 1000 volts these processes are less probable than those in which the lattice does not take up momentum. Consequently Baroody's theory of secondary emission might be essentially correct for moderately high primary energies. The theory can be extended by taking into account the motion of the lattice electrons. The probability of an energy loss between ΔE and $\Delta E+dE$, the probability of making a secondary electron with an energy between E and $E+dE$ and the total energy loss per unit path length are calculated.

* This work was supported by Signal Corps contract.
† Now at the University of British Columbia, Vancouver, Canada.
[1] E. M. Baroody, Phys. Rev. **78**, 780 (1950).
[2] D. E. Wooldridge, Phys. Rev. **56**, 562 (1939).

D7. Secondary Electron Emission under High Energy Electron Bombardment.* B. L. Miller and W. C. Porter, *Bartol Research Foundation.*—Using a linear accelerator providing pulsed electrons up to 1.5 Mev, the secondary emission of various metals has been investigated over the range 20 kilovolts to 1.5 Mev. The metals studied have been gold, tungsten, silver, stainless steel, and beryllium and the work is continuing with other metals and a few insulating materials. Comparison with the work of Trump and Van de Graaff[1] over the region 50 to 300 kilovolts, and with results obtained at the Laboratory of Nuclear Science and Engineering, M.I.T.[2] in the region 1.0 to 1.5 Mev shows qualitative agreement.

* Assisted by the joint program of the ONR and AEC.
[1] J. G. Trump and R. J. Van de Graaff, Phys. Rev. **75**, 44 (1949).
[2] Progress Report, January 1, 1951.

D8. Back-Scattering of Low Energy Electrons. E. J. Sternglass, S. C. Frey, and F. H. Grannis, *U. S. Naval Ordnance Laboratory.*—The energy dependence of electron back-scattering has been studied for carbon, copper, and molybdenum in the energy range from 0.2 to 2 kv which is of importance in the theory of secondary emission.[1] The undesirable effects of multiple scattering and secondary emission from the collector were greatly reduced by the use of spherical geometry and a suppressor screen of 68 percent open area coated with Aquadag. Appreciable backscattering was found to exist down to the lowest energies for all three elements, contrary to the expectations of some authors.[2] For an amorphous carbon film, the relative number of backscattered electrons of energy greater than 50 ev, η, was found to be essentially constant and equal to 0.07 ± 0.01. Similarly, the value for copper was found to be equal to 0.26 ± 0.02 over the whole range, in agreement with the value of 0.27 obtained by Gimpel and Richardson at 10 volts,[3] and 0.28 at 2 kv given by Palluel. For molybdenum, however, an increase of η from a value of 0.14 at 150 volts to 0.26 ± 0.02 at 1850 volts was observed. Since the data of Palluel show an increase of η with energy above 2 kv for all elements of $Z\gtrsim30$, it appears that screening and binding energy considerations play an important role in the back-scattering for the heavier elements.

[1] E. J. Sternglass, Phys. Rev. **80**, 925 (1950); thesis, Cornell University, February, 1951.
[2] P. Palluel, Compt. rend. **224**, 1492, 1551 (1947).
[3] I. Gimpel and O. Richardson, Proc. Roy. Soc. (London) **A182**, 17 (1943).

J. G. Buck presiding

Field Emission

E1. Use of Field Emission in the Study of the Absorption of Ba on W. J. A. Becker, *Bell Telephone Laboratories.*—The field emission electron microscope permits on to measure the intensity of emission from all crystallographic planes of W as functions of θ, T, and F: θ=fraction of surface covered with Ba, T=treatment temperature, F=applied field. From this one deduces how the Ba distributes itself over the surface, the ease with which it migrates on various planes, and the factors that determine evaporation rates. For θ from 0 to 1.0, the emission comes largely from Ba clusters with diameters ranging from 40 to 200A. These clusters form at or near 300°K and disappear at temperatures ranging from 370°K for 110 planes to 1000°K on 100 planes. Before the clusters disappear they are violently agitated. Above 800°K, Ba migrates quite freely everywhere. Between 1050 and 1600°K Ba evaporates; the rate depends on θ, T, and the plane. The ease of migration and the planes on which the Ba tends to congregate can be altered by the applied field. For $\theta>1$ migration is even more pronounced; crystallites with diameters from 200 to 600A form and grow outward from the surface. They thus produce high local fields and hence local enhanced emission.

E2. Field Emission; A Comparison Between Theory and Experiment Including Pulsed Emission at Large Densities.* W. P. Dyke and J. K. Trolan, *Linfield College.*—Field emission from the hemispherical single crystal tungsten emitter was obtained with pulse techniques. The empirical current-voltage relationship was confirmed in the extended range of currents from 10^{-1} to 10^{-10} ampere. The distribution of current density over the emitting surface was obtained from the densitometric analysis of the simultaneous photograph of the emission pattern when the discharge took place in a modified Müller projection tube. Experimental current densities were in good agreement with those predicted by the wave-mechanical theory when Nichol's values for the work functions of several crystal directions for clean tungsten were used. Average values of the current density, obtained from the ratio of fractional current and fractional area in the constant field region,

agreed with values predicted by the theory when the average value of the work function, $\phi = 4.5$ ev was used. Current densities of the order of 10^{-1} ampere/cm^2 were observed. The experimental error was largely determined by the exponent $\phi^{\frac{1}{2}}/F$, where F is the electric field, and this exponent was known within 16 percent for direct currents and 20 percent for pulse measurements. Vacuum arcs occurred at current densities greater than those reported.

* This work was supported by the ONR and by Research Corporation. A part of the data analysis was supported by the Microwave Laboratory of the University of California at Berkeley.

E3. Ejection of Electrons from Molybdenum by Helium Ions. H. D. HAGSTRUM, *Bell Telephone Laboratories.*—The ejection of electrons from molybdenum by the ions He^+, He^{++}, and He_2^+ has been studied. Ion kinetic energies at the target surface have been varied from 10 to 1000 electron volts. The value of the parameter γ_i (number of electrons ejected per positive ion) for each ion is found to be remarkably insensitive to the kinetic energy of the ion. γ_i (He^+) is of order 0.3, γ_i (He^{++}) of order 0.95 for clean Mo in this range. The distribution in energy of ejected electrons has also been measured. Each of these data has been obtained for clean Mo and for Mo having a monolayer of gas absorbed upon it. Variation of γ_i with formation of the monolayer has been observed. Discussion of this work in the light of other experimental work and the theory of the ejection mechanism will be presented.

E4. An Interpretation of Periodic Deviation Data for Highly Refractory Metals. E. A. COOMES, *University of Notre Dame.*—Thermionic periodic deviation data for tantalum,[1] tungsten,[2] and molybdenum[3] filaments, with clean surfaces and with monomolecular contaminations have been examined in terms of the λ- and μ-coefficients of Herring and Nichols.[4] There is fair agreement between experiment and theory for λ for the box model with a mirror image barrier. This is interpreted as indicating that a monomolecular film, either electropositive or electronegative, did not introduce a peak in the potential barrier near the surface of the composite emitter. The main effect was to lower the work function and at the same time maintain an approximate mirror image configuration. The data from these experiments do not seem to substantiate the μ-coefficient for the box model.

[1] Munick, LaBerge, and Coomes, Phys. Rev. **80**, 887 (1950).
[2] A. L. Houde, C. S. B., Ph.D. thesis, University of Notre Dame.
[3] E. G. Brock, Ph.D. thesis, University of Notre Dame.
[4] C. Herring and M. H. Nicols, Revs. Modern Phys. **21**, 185 (1949).

E5. Arcing at Electrical Contacts on Closure. Part III. Development of the Arc.* L. H. GERMER AND J. L. SMITH, *Bell Telephone Laboratories.*—When two electrodes approach each other with a small difference of potential between them an arc may be struck before they touch. This arc is initiated by field emission electrons. After the first electrons there is a lapse of time, ranging from less than 0.002 to about 0.05×10^{-6} sec, during which the arc is becoming established. For this transition time the potential between the electrodes decreases from the applied voltage to a final steady value arc voltage which is independent of current and characteristic of the metal of the electrodes. The transition depends in striking fashion upon the cleanness of the electrode surfaces but there are also erratic variations from one experiment to another; the decrease in potential is, furthermore, sometimes at a fairly steady rate but in other experiments there are irregular upward fluctuations. Oscilloscopic study of this variation of potential seems to furnish the best experimental method for investigation of the way in which an arc develops. The oscilloscopic equipment gives satisfactory photographic reproductions of single closures with a time resolution of about 0.001×10^{-6} sec. The equipment is of conventional design with only moderate modifications.

* Parts I and II, J. Appl. Phys. **22**, 955 (1951) and **22**, 1133 (1951).

E6. The Space Charge Suppression of Flicker Effect.* A VAN DER ZIEL, *University of Minnesota.*—Schottky's formula[1] for the Flicker noise in a diode under space charge conditions may be written:

$$\langle i_a^2 \rangle = (g_m V_T / I_a)^2 \langle i_s^2 \rangle. \tag{1}$$

In this equation $\langle i_s^2 \rangle$ is the fluctuation in the emission current I_s, $\langle i_a^2 \rangle$ the fluctuation in the anode current I_a, g_m is the transconductance and $V_T = kT_c/e$, where T_c is the cathode temperature. Schottky's comparison of the space-charge suppression of Flicker noise *versus* shot noise is in error, however; the corresponding formula for shot noise is

$$\langle i_a^2 \rangle = 2\theta (g_m V_T / I_a) \langle i_s^2 \rangle, \tag{2}$$

where $\theta = 3(1 - \pi/4)$. It can be shown from (1) that in those space charge limited triodes, in which the Flicker noise is due to a fluctuation in work function, the equivalent Flicker noise emf in the grid circuit of the tube is nearly equal to the fluctuation in work function. Consequently the best way of obtaining information about the fluctuating work function consists in measuring the noise resistance of space charge limited triodes as a function of frequency.

* This work was supported by Signal Corps contract.
[1] W. Schottky, Physica **4**, 175 (1937).

E7. Cold Emission of Electrons in Spark Gaps. F. LLEWELLYN JONES, *University of Wales.*—Pre-breakdown electron production, e.g., in spark gaps, is related to the time lag of sparking, and estimates of electron emission from the cathode can be made from analysis of time-lag data. In this way emission from various metal surfaces: Al, Ni, W, Fe, Cu, was measured in different gas atmospheres: dry and moist air and H_2. It was found that emission $i \simeq 10^6$ electron/sec were obtainable for most metals in the atmosphere under microscopic fields $F \simeq 10^8$ v/cm. Further, the rate of emission depended not only on the nature of the metal, but also upon its surface state, e.g., grinding a surface could increase i bv a factor $\sim 10^6$ compared with that from a polished surface. The relation between i and F was found to be $\ln(i/F^2) = A/F + B$, indicating that the extraction mechanism was a field process. The technique employed allowed effects of electron attachment and recombination to be demonstrated. In moist air, for example, post-breakdown ionization in a spark was removed after about 10^{-2} sec, while in dry air ions remained in the gap (0.025 cc) for times as long as 0.1 sec.

E8. Observations with the Field Emission Microscope. E. W. MÜLLER, *Kaiser Wilhelm Institut für Chemie.*

THE AMERICAN INSTITUTE OF PHYSICS

[illegible] 19

57 East 55 Street, New York 22, N.Y.

MEMBER SOCIETIES:

American Physical Society
Edward L. Nichols, Secretary, Columbia University, New York 27, New York

Optical Society of America
Arthur C. Hardy, Secretary, Massachusetts Institute of Technology, Cambridge 39, Massachusetts

Acoustical Society of America
Wallace Waterfall, Secretary, 57 East 55 Street, New York 22, New York

Society of Rheology
G. J. Dienes, Secretary, Brookhaven National Laboratory, Upton, Long Island, New York

American Association of Physics Teachers
R. F. Paton, Secretary, University of Illinois, Urbana, Illinois

AFFILIATED [illegible]:

American Crystallographic Association
Electron Microscopy Society of America
Physical Society of Pittsburgh
Physics Club of Chicago
Physics Club of Philadelphia
Cleveland Physics Society
Physics Club of the Lehigh Valley

ASSOCIATES:

Over seventy-five leading industrial enterprises and research institutions which contribute to the support of the Institute.

JOURNALS:

The Physical Review
S. A. Goudsmit, *Editor*, Brookhaven National Laboratory, Upton, Long Island, New York

Reviews of Modern Physics
J. W. Buchta, *Editor*, University of Minnesota, Minneapolis, Minnesota

Journal of the Optical Society of America
Wallace R. Brode, *Editor*, National Bureau of Standards, Washington 25, D.C.

The Journal of the Acoustical Society of America
Floyd R. Watson, *Editor*, 57 East 55th Street, New York 22, New York

American Journal of Physics
Thomas H. Benton, *Editor*, Michigan State College, East Lansing, Michigan

The Review of Scientific Instruments
Gaylord P. Harnwell, *Editor*, Research Laboratory of Physics, University of Pennsylvania, Philadelphia, Pennsylvania

The Journal of Chemical Physics
Joseph E. Mayer, *Editor*, Institute for Nuclear Studies, University of Chicago, Chicago 37, Illinois

Journal of Applied Physics
Ellen Swallow Richardson, *Editor*, Case Institute of Technology, Cleveland 6, Ohio

Physics Today
Gaylord P. Harnwell, *Editorial Director*; Melvin Lax, *Managing Editor*; American Institute of Physics, 57 East 55 Street, New York 22, New York

OFFICERS

Gaylord P. Harnwell, *President*
Massachusetts Institute of Technology

George B. Pegram, *Treasurer*
Columbia University

GOVERNING BOARD

J. W. Beams
C. Paul Boner
Wallace R. Brode
W. F. Faust, Jr.
S. A. Goudsmit
G. P. Harnwell
Gordon E. Hastings
F. V. Hunt
Ernest Kanzelmeyer
Paul Kirkpatrick
F. W. Loomis
J. E. Orchard
George B. Pegram
I. I. Rabi
Duane Roller
Harold K. Schilling
Frederick Seitz
Francis G. Slack
John C. Slater

Committee on Physics
Melvin Lax

ADMINISTRATIVE STAFF

Henry A. Barton, *Executive Secretary*
Wallace Waterfall, *Executive Secretary*
Theodora Vogt, *Advertising Manager*
Ruth F. Bryant, *Publication Manager*
Betty H. Goodwin, *Assistant Secretary*
Eileen B. Nugent, *Circulation Manager*
Kathryn Sellers, *Assistant Treasurer*

THE American Institute of Physics was founded in 1931 as a federation of leading societies in the field of Physics. It combines into one operating unit those functions on behalf of physics which can best be developed by the societies jointly. Its purpose is the advancement and diffusion of physics and its applications to human welfare. To this end it encourages for itself or the Societies participating journals listed on this page; promotes unity and effectiveness of effort among all organizations interested in physics; renders various direct services to physicists, to science, to the public; and cooperates with governmental agencies, foundations, associations, educational institutions, and industrial laboratories in such manner as to further the opportunities and facilities for the [illegible] and constructive [illegible]